Biosynthesis and transfer genes of 6-desoxyhexoses in *Saccharopolyspora erythraea* and in *Streptomyces antibioticus* and their use.

The present invention describes the genes involved in the biosynthesis and transfer of 6-desoxyhexoses in *Saccharopolyspora erythraea* and their use in the production of analogues of erythromycin by genetic manipulation.

Erythromycin A is a clinically important macrolide antibiotic produced by the gram-positive bacterium *Sac. erythraea*. The biosynthesis genes of erythromycin are organised in a cluster of *ery* genes which also includes the erythromycin auto-resistance gene *erm*E.

The *ery* cluster contains the three major genes *ery*AI, *ery*AII and *ery*AIII (locus *ery*A) coding for three polypeptides comprising the polyketide synthetase (called PKS) flanked by two regions comprising the genes involved in the subsequent stages of the conversion of the lactone nucleus (6-desoxyerythronolide B) into erythromycin A.

During the erythromycin A biosynthesis process represented in Figure 1, the biosynthesis of the 6-desoxyhexoses comprises all the enzymatic reactions leading from glucose-1-phosphate to the final activated sugar dTDP-L-mycarose or dTDP-D-desosamine. The dTDP-L-mycarose or dTDP-D-desosamine thus produced are then used as substrates for the transfer of the two desoxyhexoses onto the lactone nucleus. The formation of erythromycin requires the attachment of the mycarose via the hydroxyl in position C-3 of the lactone nucleus and the attachment of the desosamine via the hydroxyl in position C-5. All the *ery*B genes involved in the biosynthesis or the transfer of the mycarose and all the *ery*C genes involved in the biosynthesis or transfer of the desosamine have not yet been clearly identified.

The *ery* cluster with a length of 56 kb comprises 21 open reading frames (ORFs) the numbering of which has been established by Haydock et al. (1991) and Donadio et al. (1993). The *ery*A locus comprises the ORFs 10, 11 and 12.

Early work on gene interruption or replacement in the left part of the *ery* cluster allowed a first identification of the *ery*CI gene (ORF1) (Dhillon et al., 1989), then of the *ery*BI gene (ORF2), of the *ery*H locus (ORFs 3, 4 and 5) the inactivation of which leads to the production of 6-desoxyerythronolide B, of a *ery*BII locus (ORFs 7 and 8) and the *ery*CII gene (Weber et al., 1990).

Among the enzymatic activities involved in the subsequent modifications of the lactone nucleus the *ery*F gene (ORF4) responsible for the hydroxylation in position C6 (Weber et al., 1991) and the *ery*K gene (ORF20) responsible for the hydroxylation in C12 (Stassi et al., 1993) have been identified. Moreover, the *ery*G gene (ORF6) responsible for the O-methylation of mycarose to cladinose (position 3"OH) has been identified (Weber et al., 1989). Erythromycin A is thus formed via erythromycin B or erythromycin C from erythromycin D according to the proposed diagram (Figure 1).

The functional characterization of the *ery*B and *ery*C genes situated on the right part of the *ery* cluster (ORFs 13 to 19) has not yet been precisely established, despite the fragmentary information reported in various journals(Donadio et al., 1993; Liu and Thorson, 1994; Katz and Donadio, 1995).

Due to the commercial interest in macrolide antibiotics, the obtaining of new derivatives, in particular of analogues of erythromycin having advantageous properties, is keenly sought. Modifications may be desired in the aglycone part (macrolactone) or/and in its secondary hydroxylation as well as in the sugar part (cladinose and/or desosamine) of erythromycin.

Current methods such as chemical modifications are difficult and limited vis-à-vis the type of product that can be obtained from erythromycin. For example, Sakakibara et al. (1984) review chemical modifications carried out from erythromycin A or B, both in the sugar part and the macrolactone.

Modifications of the macrolactone of erythromycin A by genetic manipulation of the microorganism *Sac. erythraea* have been described in the International Patent Application

WO 93/13663 as well the obtaining of novel polyketide molecules by specific genetic alterations of the *ery*A locus of the chromosome coding for PKS. For example 7-hydroxyerythromycin A, 6-desoxy-7-hydroxyerythromycin A or 3-oxo-3-desoxy-5-desoaminyl-erythronolide A have been obtained in this way.

The present invention relates to the functional characterization of ten genes of *Sac. erythraea* involved in the biosynthesis or the attachment of mycarose and of desosamine (*ery*BII, *ery*CIII and *ery*CII situated downstream from the *ery*A locus and *ery*BIV, *ery*BV, *ery*CVI, *ery*BVI, *ery*CIV, *ery*CV and *ery*BVII situated upstream), their use in the production of analogues of erythromycin as well as a preparation process for them.

Therefore a subject of the present invention is an isolated single- or double-stranded DNA sequence, represented in Figure 2 (direct and complementary sequence of SEQ ID No 1) corresponding to the *ery*G-*ery*AIII region of the cluster of erythromycin biosynthesis genes and a particular subject is an above DNA sequence comprising:

- the *ery*BII sequence corresponding to ORF7 (complementary sequence of SEQ ID No.1 from nucleotide 48 to nucleotide 1046) and coding for a dTDP-4-keto-L-6-desoxyhexose 2,3-reductase,
- the *ery*CIII sequence corresponding to ORF8 (complementary sequence of SEQ ID No.1 from nucleotide 1046 to nucleotide 2308) and coding for a desosaminyltransferase and
- the *ery*CII sequence corresponding to ORF9 (complementary sequence of SEQ ID No.1 from nucleotide 2322 to nucleotide 3404) and coding for a dTDP-4-keto-D-6-desoxyhexose 3,4-isomerase.

The above DNA sequence shown in Figure 2 is a genomic DNA sequence which can be obtained for example by sub-cloning restriction fragments of a genomic DNA fragment of *Sac. erythraea*, according to operating conditions a detailed description of which is given hereafter.

A more particular subject of the invention is an isolated DNA sequence represented in Figure 2 chosen from the

*ery*BII sequence corresponding to ORF7 (complementary sequence of SEQ ID No.1 from nucleotide 48 to nucleotide 1046), the *ery*CIII sequence corresponding to ORF8 (complementary sequence of SEQ ID No.1 from nucleotide 1046 to nucleotide 2308) or the *ery*CII sequence corresponding to ORF9 (complementary sequence of SEQ ID No.1 from nucleotide 2322 to nucleotide 3404) and the sequences which hybridize and/or have significant homologies with this sequence or fragments of it and having the same function.

A quite particular subject of the invention is the *ery*CIII isolated DNA sequence represented in Figure 2 corresponding to ORF8 (complementary sequence of SEQ ID No.1 from nucleotide 1046 to nucleotide 2308 = complementary sequence of SEQ ID No.4) and coding for a desosaminyl-transferase.

The *ery*BII sequence corresponding to ORF7 codes for a polypeptide having 333 amino acids (sequence of SEQ ID No.2), the *ery*CIII sequence corresponding to ORF8 codes for a polypeptide having 421 amino acids (sequence of SEQ ID No.5) and the *ery*CII sequence corresponding to ORF9 codes for a polypeptide having 361 amino acids (sequence of SEQ ID No.3).

The respective enzymatic activities indicated above has been demostrated by the introduction of an internal deletion to the corresponding gene as illustrated hereafter in the experimental part.

Sequences which hybridize and having the same function are taken to include DNA sequences which hybridize with one of the above DNA sequences under standard conditions of high or medium stringency described by Sambrook et al. (1989) and which code for a protein having the same enzymatic function. The term same enzymatic function, is taken to mean a given enzymatic activity on substrates of a similar nature, for example a dTDP-6-desoxyhexose or a bare or glycosylated macrolactone. Conditions of high stringency comprise for example a hybridization at 65°C for 18 hours in a 5 x SSPE, 10 x Denhardt, 100 µg/ml DNAss, 1 % SDS solution, followed by 2 washings for 20 minutes with a 2 x SSC, 0.05 % SDS solution at 65°C followed by a last washing for 45 minutes in a

0.1 x SSC, 0.1 % SDS solution at 65°C. Conditions of medium stringency comprise for example a last washing for 20 minutes in a 0.2 x SSC, 0.1 % SDS solution at 65°C.

Sequences which display significant homologies and having the same function are taken to include sequences having an identity of nucleotide sequence of at least 60 % with one of the above DNA sequences and which code for a protein having the same enzymatic function.

A subject of the invention is also a polypeptide coded by one of the above DNA sequences and a particular subject is a polypeptide corresponding to an ORF represented in Figure 2, chosen from ORF7 (having the sequence of SEQ ID No.2), ORF8 (having the sequence of SEQ ID No.5) or ORF9 (having the sequence of SEQ ID No.3) and the analogues of this polypeptide.

Analogues are taken to include the peptides having an amino acid sequence modified by substitution, deletion or addition of one or more amino acids inasmuch as these products retain the same enzymatic function. The modified sequences can be prepared, for example, by using the site-directed mutagenesis technique known to a person skilled in the art.

A more particular subject of the invention is the polypeptide corresponding to ORF 8 represented in Figure 2 (having the sequence of SEQ ID No.5) and having a desosaminyltransferase activity, called EryCIII.

The invention describes an EryCIII recombinant protein of *Sac. erythraea* obtained by expression in a host cell according to known genetic engineering and cell culture methods.

The obtaining of the purified recombinant protein has allowed confirmation of the characterization of the glycosyltranferase function associated with the product of the *ery*CIII gene in an in-vitro test which demonstrates the transfer of the active sugar dTDP-D-desosamine onto the lactone nucleus.

A subject of the invention is also thymidine 5'-(trihydrogen diphosphate),P'-[3.4,6-tridesoxy-3-(dimethylamino)-

D-.xylo.-hexopyranosyl] ester (dTDP-D-desosamine) and the addition salts with the bases, a preparation example of which is described hereafter in the experimental part.

A subject of the invention is also an isolated DNA sequence represented in Figure 3 (sequence of SEQ ID No.6) corresponding to the *ery*AI-*ery*K region of the cluster of erythromycin biosynthesis genes and a particular subject is an above DNA sequence comprising:

- the *ery*BIV sequence corresponding to ORF13 (sequence of SEQ ID No.6 from nucleotide 242 to nucleotide 1207) and coding for a dTDP-4-ceto-L-6-desoxyhexose 4-reductase,
- the *ery*BV sequence corresponding to ORF14 (sequence of SEQ ID No.6 from nucleotide 1210 to nucleotide 2454) and coding for a mycarosyltransferase,
- the *ery*CVI sequence corresponding to ORF15 (sequence of SEQ ID No.6 from nucleotide 2510 to nucleotide 3220) and coding for a dTDP-D-6-desoxyhexose 3-N-methyltransferase,
- the *ery*BVI sequence corresponding to ORF16 (sequence of SEQ ID No.6 from nucleotide 3308 to nucleotide 4837) and coding for a dTDP-4-ceto-L-6-desoxyhexose 2,3-deshydratase,
- the *ery*CIV sequence corresponding to ORF17 (sequence of SEQ ID No.6 from nucleotide 4837 to nucleotide 6039) and coding for a dTDP-D-6-desoxyhexose 3,4-deshydratase,
- the *ery*CV sequence corresponding to ORF18 (sequence of SEQ ID No.6 from nucleotide 6080 to nucleotide 7546) and coding for a dTDP-D-4,6-didesoxyhexose 3,4-reductase and
- the *ery*BVII sequence corresponding to ORF19 (sequence of SEQ ID No.6 from nucleotide 7578 to nucleotide 8156) and coding for a dTDP-4-ceto-D-6-desoxyhexose 3,5 epimerase.

The above DNA sequence shown in Figure 3 is a genomic DNA sequence which can be obtained, for example, by sub-cloning restriction fragments of cosmids containing a genomic DNA library of *Sac. erythraea*, according to the operating conditions a detailed description of which is given hereafter.

A more particular subject of the invention is an isolated DNA sequence represented in Figure 3 chosen from the *ery*BIV sequence corresponding to ORF13 (sequence of SEQ ID

No.6 from nucleotide 242 to nucleotide 1207), the *ery*BV sequence corresponding to ORF14 (sequence of SEQ ID No.6 from nucleotide 1210 to nucleotide 2454), the *ery*CVI sequence corresponding to ORF15 (sequence of SEQ ID No.6 from nucleotide 2510 to nucleotide 3220), the *ery*BVI sequence corresponding to ORF16 (sequence of SEQ ID No.6 from nucleotide 3308 to nucleotide 4837), the *ery*CIV sequence corresponding to ORF17 (sequence of SEQ ID No.6 from nucleotide 4837 to nucleotide 6039), the *ery*CV sequence corresponding to ORF18 (sequence of SEQ ID No.6 from nucleotide 6080 to nucleotide 7546) or the *ery*BVII sequence corresponding to ORF19 (sequence of SEQ ID No.6 from nucleotide 7578 to nucleotide 8156) and the sequences which hybridize and/or display significant homologies with this sequence or fragments of the latter and having the same function.

A quite particular subject of the invention is the *ery*BV isolated DNA sequence represented in Figure 3 corresponding to ORF14 (sequence of SEQ ID No.6 from nucleotide 1210 to nucleotide 2454) and coding for a mycarosyltransferase.

The *ery*BIV sequence corresponding to ORF13 codes for a polypeptide having 322 amino acids (SEQ ID No.7), the *ery*BV sequence corresponding to ORF14 codes for a polypeptide having 415 amino acids (SEQ ID No.8), the *ery*CVI sequence corresponding to ORF15 codes for a polypeptide having 237 amino acids (SEQ ID No.9), the *ery*BVI sequence corresponding to ORF16 codes for a polypeptide having 510 amino acids (SEQ ID No.10), the *ery*CIV sequence corresponding to ORF17 codes for a polypeptide having 401 amino acids (SEQ ID No.14), the *ery*CV sequence corresponding to ORF18 codes for a polypeptide having 489 amino acids (SEQ ID No.11) and the *ery*BVII sequence corresponding to ORF19 codes for a polypeptide having 193 amino acids (SEQ ID No.12).

The respective enzymatic activities indicated above have been demonstrated by the introduction of an internal deletion to the corresponding gene as illustrated hereafter in the experimental part.

The DNA sequences which hybridize as well as the DNA

sequences which display significant homologies and having the same function have the same meaning as that indicated previously.

A subject of the invention is also a polypeptide coded by one of the above DNA sequences and a particular subject is a polypeptide corresponding to an ORF represented in Figure 3, chosen from ORF13 (having the sequence of SEQ ID No.7), ORF14 (having the sequence of SEQ ID No.8), ORF15 (having the sequence of SEQ ID No.9), ORF16 (having the sequence of SEQ ID No.10), ORF17 (having the sequence of SEQ ID No.14), ORF18 (having the sequence of SEQ ID No.11) or ORF19 (having the sequence of SEQ ID No.12) and the analogues of this peptide.

The analogues of the polypeptide have the same meaning as that indicated previously.

A more particular subject of the invention is the polypeptide corresponding to ORF14 represented in Figure 3 (having the sequence of SEQ ID No.8) and having a mycarosyltransferase activity, called EryBV.

Knowledge of each *ery*B or *ery*C DNA sequence of the invention indicated above and shown in Figure 2 or Figure 3 allows the present invention to be reproduced for example by known methods of chemical synthesis or by screening a genomic library using oligonucleotide synthesis probes by known techniques of hybridization or by PCR amplification.

The polypeptides of the invention can be obtained by known methods, for example by chemical synthesis or by the recombinant DNAmethod by expression in a procaryotic or eucaryotic host cell.

Another subject of the invention relates to the use of at least one of the DNA sequences chosen from the sequences *ery*BII (complementary sequence of SEQ ID No.1 from nucleotide 48 to nucleotide 1046), *ery*CIII (complementary sequence of SEQ ID No.1 from nucleotide 1046 to nucleotide 2308) or *ery*CII (complementary sequence of SEQ ID No.1 from nucleotide 2322 to nucleotide 3404) represented in Figure 2, *ery*BIV (sequence of SEQ ID No.6 from nucleotide 242 to nucleotide 1207), *ery*BV (sequence of SEQ ID No.6 from nucleotide 1210 to nucleotide 2454), *ery*CVI (sequence of SEQ ID No.6 from

nucleotide 2510 to nucleotide 3220), *ery*BVI (sequence of SEQ ID No.6 from nucleotide 3308 to nucleotide 4837), *ery*CIV (sequence of SEQ ID No.6 from nucleotide 4837 to nucleotide 6039), *ery*CV (sequence of SEQ ID No.6 from nucleotide 6080 to nucleotide 7546) or *ery*BVII (sequence of SEQ ID No.6 from nucleotide 7578 to nucleotide 8156) represented in Figure 3, to synthesize hybrid secondary metabolites in *Sac. erythraea*.

Hybrid secondary metabolites are taken to mean either analogues of erythromycin, i.e. derivatives of erythromycin having one or more modifications affecting the sugar part and having an antibiotic activity, or precursors of erythromycin such as 6-desoxyerythronolide B or erythronolide B to which one or more sugar residues modified or not and not having an antibiotic activity are attached. The modified sugar residue can be, for example, 4-keto-L-mycarose.

The synthesis of hybrid secondary metabolites in *Sac. erythraea* by the use of *ery*B or *ery*C DNA sequences of the invention can be carried out, for example, by the inactivation of one or more *ery*B or *ery*C genes above and the introduction of one or more exogenous genes or their derivatives obtained for example by mutagenesis, having nucleotide sequences coding for the enzymes having the same function in strains producing other macrolides, for example tylosin, picromycin or methymycin. In particular, the introduction of exogenous genes can be carried out by the integration of a DNA sequence obtained according to the "DNA shuffling" method (Stemmer, 1994) or by the construction of a chimeric DNA sequence, for example from an *ery*B or *ery*C sequence of the invention intervening in the transfer of a sugar residue, for example the *ery*CIII or *ery*BV sequence, and homologous genes isolated from strains producing macrolides, for example *Streptomyces fradiae, Streptomyces olivaceus, Streptomyces venezuelae or Streptomyces antibioticus*.

The invention also relates to the use of at least one of the DNA sequences chosen from the sequences *ery*BII (complementary sequence of SEQ ID No.1 from nucleotide 48 to nucleotide 1046), *ery*CIII (complementary sequence of SEQ ID No.1 from nucleotide 1046 to nucleotide 2308) or *ery*CII

(complementary sequence of SEQ ID No.1 from nucleotide 2322 to nucleotide 3404) represented in Figure 2, *ery*BIV (sequence of SEQ ID No.6 from nucleotide 242 to nucleotide 1207), *ery*BV (sequence of SEQ ID No.6 from nucleotide 1210 to nucleotide 2454), *ery*CVI (sequence of SEQ ID No.6 from nucleotide 2510 to nucleotide 3220), *ery*BVI (sequence of SEQ ID No.6 from nucleotide 3308 to nucleotide 4837), *ery*CIV (sequence of SEQ ID No.6 from nucleotide 4837 to nucleotide 6039), *ery*CV (sequence of SEQ ID No.6 from nucleotide 6080 to nucleotide 7546) or *ery*BVII (sequence of SEQ ID No.6 from nucleotide 7578 to nucleotide 8156) represented in Figure 3 or of a fragment of this sequence, as hybridization probes.

The *ery*B or *ery*C DNA sequences of the invention can be used to constitute hybridization probes of at least 19 nucleotides, allowing homologous genes to be isolated in strains producing macrolides by using the standard methods of hybridization of nucleic acids immobilized on filters or of PCR amplification, according to the conditions described by Sambrook et al. (1989).

The invention particularly relates to the use of the *ery*CIII DNA sequence represented in Figure 2 (complementary sequence of SEQ ID No.1 from nucleotide 1046 to nucleotide 2308 = complementary sequence of SEQ ID No.4) as a hybridization probe for isolating homologous genes responsible for the glycosylation of the macrolactone in a productive strain of macrolide.

The invention more particularly relates to the above use, in which the homologous genes are the oleandomycin biosynthesis genes in *S. antibioticus*.

The invention describes, by way of an example, the use of the sequence of the *ery*CIII gene as a hybridization probe for isolating homologous genes in a productive strain of oleandomycin. The *ery*CIII probe used has allowed the isolation of the *ole*G1 and *ole*G2 genes coding for glycosyltransferases in *S. antibioticus* involved in the transfer of desosamine and oleandrose onto the lactone nucleus.

The functional characterization of the *ole*G1 and *ole*G2

genes has allowed the organisation of the right part of the cluster of oleandomycin biosynthesis genes in *S. antibioticus* to be defined.

Therefore a subject of the invention is an isolated DNA sequence represented in Figure 22 (sequence of SEQ ID No.15) corresponding to a region of the cluster of oleandomycin biosynthesis genes comprising:

- the sequence corresponding to ORF *ole*P1 from nucleotide 184 to nucleotide 1386,
- the sequence corresponding to ORF *ole*G1 from nucleotide 1437 to nucleotide 2714 coding for a glycosyltransferase activity,
- the sequence corresponding to ORF *ole*G2 from nucleotide 2722 to nucleotide 3999 coding for a glycosyltransferase activity,
- the sequence corresponding to ORF *ole*M from nucleotide 3992 to nucleotide 4720 (= sequence of SEQ ID No.20) and
- the sequence corresponding to ORF *ole*Y from nucleotide 4810 to nucleotide 5967.

The above DNA sequence shown in Figure 22 (sequence of SEQ ID No.15) is a genomic DNA sequence which can be obtained for example from a cosmid covering the right part of the cluster of oleandomycin biosynthesis genes by hybridization with an *ery*CIII probe, according to the operating conditions a detailed description of which is given hereafter.

A more particular subject of the invention is an isolated DNA sequence represented in Figure 22 chosen from the sequence corresponding to ORF *ole*G1 (sequence of SEQ ID No.15 from nucleotide 1437 to nucleotide 2714 coding for a glycosyltransferase activity and the sequence corresponding to ORF *ole*G2 (sequence of SEQ ID No.15 from nucleotide 2722 to nucleotide 3999) coding for a glycosyltransferase activity.

A quite particular subject of the invention is an isolated DNA sequence above corresponding to ORF *ole*G1 (sequence of SEQ ID No.15 from nucleotide 1437 to nucleotide 2714) coding for a desosaminyltransferase activity, as well as an isolated DNA sequence above corresponding to ORF *ole*G2

(sequence of SEQ ID No.15 from nucleotide 2722 to nucleotide 3999) coding for an oleandrosyltransferase activity.

The sequence corresponding to ORF *ole*G1 codes for a polypeptide having 426 amino acids (sequence of SEQ ID No.17) and the sequence corresponding to ORF *ole*G2 codes for a polypeptide having 426 amino acids (sequence of SEQ ID No.18).

The respective enzymatic activities indicated above have been demonstrated by alteration of the corresponding gene as illustrated hereafter in the experimental part.

A subject of the invention is also the polypeptide coded by the DNA sequence corresponding to ORF *ole*G1 and having a desosaminyltransferase activity (sequence of SEQ ID No.17) and the polypeptide coded by the DNA sequence corresponding to ORF *ole*G2 and having an oleandrosyltransferase activity (sequence of SEQ ID No.18).

The polypeptides above called OleG1 and OleG2 respectively can be obtained by the known methods indicated above.

A subject of the invention is also a process for the preparation of hybrid secondary metabolites in *Sac. erythraea* in which:

- a DNA sequence is isolated containing at least one *ery*B sequence or an *ery*C sequence of the cluster of biosynthesis genes of erythromycin represented in Figure 2 (complementary sequence of SEQ ID No.1) or in Figure 3 (sequence of SEQ ID No.6),
- a modification is created in said sequence and an altered sequence is obtained,
- the altered sequence is integrated into the chromosome of the host strain and a modified strain is obtained,
- the modified strain is cultured under conditions which allow the formation of the hybrid secondary metabolite and
- the hybrid secondary metabolite is isolated.

The modification of the DNA sequence can be carried out for example by an addition and/or by a deletion of DNA sequences of at least one nucleotide, in an *ery*B or *ery*C sequence of the invention which codes for one of the

corresponding enzymes indicated above.

The integration of the altered sequence into the host strain can be carried out for example by the homologous recombination method which can be carried out according to the diagram shown in Figure 4 and leads to the generation of chromosomal mutants of *Sac. erythraea* strains which are then cultured according to the known general methods of cell culture.

A particular subject of the invention is the above process in which the DNA sequence codes for one of the enzymes chosen from a

- dTDP-4-keto-D-6-desoxyhexose 2,3-reductase,
- desosaminyltransferase,
- dTDP-4-keto-D-6-desoxyhexose 3,4-isomerase,
- dTDP-4-keto-L-6-desoxyhexose 4-reductase,
- mycarosyltransferase,
- dTDP-D-6-desoxyhexose 3-N-methyltransferase,
- dTDP-4-keto-L-6-desoxyhexose 2,3-deshydratase,
- dTDP-D-6-desoxyhexose 3,4-deshydratase,
- dTDP-D-4,6-didesoxyhexose 3,4-reductase or
- dTDP-4-keto-D-6-desoxyhexose 3,5 epimerase.

A more particular subject of the invention is the above process in which the alteration of the sequence results in the inactivation of at least one of the enzymes indicated above.

The inactivation of at least one of the enzymes is demonstrated, on the one hand by the absence of production of erythromycin, on the other hand by the accumulation of precursors of erythromycin such as 6-desoxyerythronolide B, erythronolide B or 3-α-mycarosyl erythronolide B and/or the accumulation of hybrid secondary metabolites as defined previously in the supernatants of cultures of the corresponding modified strains.

The invention quite particularly relates to the above process in which the inactivated enzyme is a dTDP-4-keto-L-6-desoxyhexose 4-reductase or in which the inactivated enzyme is a dTDP-D-6-desoxyhexose 3,4-deshydratase or in which the enzyme is a mycarosyltransferase or in which the inactivated

enzyme is a dTDP-4-keto-L-6-desoxyhexose 2,3-reductase.

The invention also relates to the above process in which the isolated hybrid secondary metabolite is an analogue of erythromycin chosen from 4"-keto-erythromycin, 4'-hydroxy-erythromycin or 3"-C desmethyl-2",3"-ene-erythromycin or in which the isolated hybrid secondary metabolite is desosaminyl erythronolide B.

Examples of the implementation of the process according to the invention are given in the experimental part. The accumulation of hybrid secondary metabolites in modified strains of *Sac. erythraea* is also described hereafter.

The invention also relates to a modified strain of *Sac. erythraea* in which at least one of the enzymes is chosen from a

- dTDP-4-keto-L-6-desoxyhexose 2,3-reductase,
- desosaminyltransferase,
- dTDP-4-keto-D-6-desoxyhexose 3,4-isomerase,
- dTDP-4-keto-L-6-desoxyhexose 4-reductase,
- mycarosyltransferase,
- dTDP-D-6-desoxyhexose 3-N-methyltransferase,
- dTDP-4-keto-L-6-desoxyhexose 2,3-deshydratase,
- dTDP-D-6-desoxyhexose 3,4-deshydratase,
- dTDP-D-4,6-didesoxyhexose 3,4-reductase or
- dTDP-4-keto-D-6-desoxyhexose 3,5 epimerase

is inactivated and producing at least one hybrid secondary metabolite.

The invention particularly relates to the modified strain of *Sac. erythraea* BII92 in which a dTDP-4-keto-L-6-desoxyhexose 2,3-reductase is inactivated and producing 3"-C desmethyl-2",3"-ene-erythromycin C, the modified strain BIV87 of *Sac. erythraea* in which a dTDP-4-keto-L-6-desoxyhexose 4-reductase is inactivated and producing 4"-keto-erythromycin, the modified strain CIV89 of *Sac. erythraea* in which a dTDP-D-6-desoxyhexose 3,4-deshydratase is inactivated and producing 4'-hydroxyerythromycin D as well as the modified strain BV88of *Sac. erythraea* in which a mycarosyltransferase is inactivated and producing desoaminyl erythronolide B. Detailed constructions of the above strains are given

hereafter in the experimental part.

The invention also relates to a process for the preparation of precursors of oleandomycin in *S. antibioticus* in which

- an alteration is created in the sequence of the gene chosen from the DNA sequence corresponding to ORF *ole*G1 (sequence of SEQ ID No.15 from nucleotide 1437 to nucleotide 2714)
and the DNA sequence corresponding to ORF *ole*G2 (sequence of SEQ ID No.15 from nucleotide 2722 to nucleotide 3999) in the chromosome of a host strain and a modified strain is obtained,
- the modified strain is cultured under conditions allowing the accumulation of the precursors of oleandomycin and
- these precursors are isolated.

The alteration of the DNA sequence can be carried out for example by interruption of the target gene in the strain *S. antibioticus*, for example by integration of a plasmid by the homologous recombinant method and leads to the generation of chromosomal mutants of the wild strain.

The invention particularly relates to an above process in which the alteration is created in the DNA sequence corresponding to ORF *ole*G1 (sequence of SEQ ID No.15 from nucleotide 1437 to nucleotide 2714) and the result of which is at least the elimination of the desoaminyltransferase activity and the accumulation of the precursor of oleandomycin 8,8a-desoxyoleandolide.

The accumulation of a non-glycosylated precursor of oleandomycin 8,8a-desoxyoleandolide observed by interruption of the *ole*G1 gene is due to a transcriptional polar effect which inactivates the *ole*G2 gene.

An example of implementation of the above process is given hereafter in the experimental part.

Materials and general methods.

1. Bacterial strains, plasmids and growth conditions.

The *Sac. erythraea* strain used for the realisation of the invention is a spontaneous phenotypic variant called "red

variant" (Hessler et al., 1997) of the wild strain *Sac. erythraea* NRRL 2338 the growth of which is carried out in a routine manner either on solid medium R2T2 (medium R2T described by Weber et al., 1985 without peptone), R2T20 (Yamamoto et al., 1986) or M1-102 on agar (Kaneda et al., 1962), or in liquid medium TSB (Oxoid) at 30°C.

The *Streptomyces lividans* strain 1326 (John Innes Culture Collection) described by Hopwood et al. (1983), used for the preparation of plasmids devoid of replication origin of *Escherichia coli* such as pIJ702 and pIJ486, was maintained on solid medium R2YE(R5) (Hopwood et al., 1985).

Growth of the *E. coli* XL1-blue (Stratagene), JM110 (Stratagene) and DH5α.MCR (GibcoBRL) strains, used for preparations of plasmids, was carried out in a routine manner in liquid medium 2 x YT or LB or in solid medium LB on agar, as described by Sambrook et al. (1989). The *E. coli* XL1-blue strain is used for routine clonings. The JM110 strain is used for clonings where restriction sites such as *Bcl*I are used. The DH5α.MCR strain is used for the preparation of plasmids intended to be introduced into *Sac. erythraea* for an optimum transformation.

Selection of the plasmids in *E. coli* was carried out on ampicillin (Sigma) at 100 µg/ml.

The *Bacillus subtilis* ATCC 6633 or *Bacillus pumilus* ATCC 14884 strains were used as indicator strains to evaluate the production of erythromycin in biological tests using a record of bacterial sensitivity to antibiotics.

The Litmus28, pUC18 and pUC19 plasmids (New England Biolabs) were used in a routine manner for the sub-clonings. The pIJ702 vector (Katz et al., 1983) was obtained from the John Innes Institute. The pIJ486 vector (Ward et al., 1986) was obtained from C.J. Thompson (University of Basle, Switzerland). The phagmide pTZ18R was obtained from Pharmacia Biotech. The shuttle vector coli-streptomyces pUWL218 (Wehmeier, 1995) used for the chromosomal integration intp *Sac. erythraea* was obtained from W.Piepersberg (Wuppertal University, Germany).

2. DNA manipulation and sequencing.

The general methods of molecular biology used are described by Sambrook et al., 1989.

Reagents of commercial origin were used including the restriction enzymes (New England Biolabs and Boehringer Mannheim), the Klenow fragment of DNA polymerase I (Boehringer Mannheim). The "DNA ligation system" kit (Amersham) was used to carry out the ligations and the Plasmid Midi kit (Quiagen) or RPM kit (Bio101 Inc.) for purifying the plasmidic DNA.

The preparation of the DNA of the λ bacteriophage was carried out according to Ausubel et al. (1995) and the isolation of the chromosomal DNA of *Sac. erythraea* according to Hopwood et al. (1985).

The transformation of *S. lividans* and isolation of the plasmids was carried out according to Hopwood et al. (1985).

3. Preparation of erythronolide B and 3-α-mycarosyl erythronolide B.

Erythronolide B and 3-α-mycarosyl erythronolide B were purified from culture extracts of the mutant *ery*CI (clone WHB2221 described by Dhillon et al., 1989) by chromatography on aminopropyl gel (LichroprepNH2 25-40 μ, Merck) with an elution gradient by successive butyl chloride/ methylene chloride mixtures (100:0, 80:20, 50:50 and 20:80) followed by a linear elution gradient by the butyl chloride /methanol mixture varying from 99:1 to 90:10. The fractions containing the expected products are brought to dryness under vacuum then analyzed by thin-layer chromatography (TLC). Erythronolide B is then crystallized from an ethyl acetate/hexane mixture then recrystallized from ethanol. 3-α-mycarosyl erythronolide B is crystallized twice from an ethyl acetate/hexane mixture.

Mentioned media.

1. R2T2:

For 1 litre of aqueous solution: sucrose 103 g; K_2SO_4 0.25g; yeast extract 6.5 g; tryptone 5.0 g; bactoagar 22.0 g; distilled water sqf 860 ml. The solution is sterilized by autoclaving for 30 minutes at 120°C. At the time of use, the following sterile solutions are added: 20 ml of glucose at 50

%; 25 ml of Tris-HCl 1M, pH7.0; 5 ml of KH_2PO_4 at 0.5 %; 2.5 ml of NaOH 1N; 50 ml of $CaCl_2$ 1M; 50 ml of $MgCl_2.6H_2O$ 1M and 2 ml of solution of "trace elements" (Hopwood et al., 1985).

2. R2T20:

For one litre of aqueous solution: R2T2 medium containing 206 g of sucrose.

3. M1-102 (Kaneda et al., 1962):

For 1 litre of aqueous solution: glucose 5 g; commercial brown sugar 10 g; tryptone 5 g; yeast extract 2.5 g; Versene 36 mg; tap water 1000 ml; final pH adjusted to 7.0 to 7.2 with KOH. The solution is sterilized by autoclaving for 30 minutes at 120°C.

4. R2YE(R5) (Hopwood et al., 1985):

For 1 litre of aqueous solution: sucrose 103 g; K_2SO_4 0.25 g; $MgCl_2.6H_2O$ 10.12 g; casaminoacids 0.1 g; solution of "trace elements" 2 ml; yeast extract 5 g; TES 5.72 g; bactoagar 15 g; distilled water sqf 940 ml. The solution is sterilized by autoclaving for 30 minutes at 120°C. At the time of use, the following sterile solutions are added: 10 ml of 0.5 % KH_2PO_4; 20 ml of $CaCl_2$ 1M; 15 ml of L-proline at 20 %; 20 ml of glucose at 50 % and 1 ml of $CuCl_2$ 10mM.

5. 2 x TY:

For 1 litre of aqueous solution: tryptone 10 g; yeast extract 10 g; NaCl 5 g.

6. PT buffer:

For 1 litre of aqueous solution: sucrose 100 g; K_2SO_4 0.25 g; $MgCl_26H_2O$ 5.1 g; solution of "trace elements" 2 ml; distilled water sqf 875 ml. The solution is sterilized by autoclaving for 30 minutes at 120°C. At the time of use, the following sterile solutions are added: 5 ml of $CaCl_2$ and 20 ml of TES 5.3 %.

7. Sucrose-succinate (Caffrey et al., 1992):

For 1 litre of aqueous solution: sucrose 0.2 M; succinic acid 20 mM; potassium phosphate 20 mM (pH 6.6); magnesium sulphate 5 mM; potassium nitrate 100 mM; solution of "trace elements" 2 ml. The solution is sterilized by autoclaving for 30 minutes at 120°C.

The attached figures illustrate certain aspects of the

invention.

Figure 1 represents the biosynthesis route for erythromycin A.

Figure 2 represents the nucleotide sequence (direct and complementary sequence of SEQ ID No.1) of the *ery*G-*ery*AIII region of the cluster of biosynthesis genes of erythromycin comprising ORFs 7, 8 and 9 and their deduced protein sequences.

Figure 3 represents the nucleotide sequence (sequence of SEQ ID No.6) of the *ery*AI-*ery*K region of the cluster of biosynthesis genes of erythromycin comprising ORFs 13 to 19 and their deduced protein sequences.

Figure 4 represents the diagram for gene substitution by homologous recombination.

Figure 5A represents the organisation of the left part of the cluster of biosynthesis genes of erythromycin in *Sac. erythraea* the ORFs 1 to 9 of which are indicated by arrows as well as a restriction map of the plasmids pK62, pBCK1, pKB22, pBK44, pBIISB, pEco2 and pK23, generated from genomic clone λSE5.5. (Abbreviations of the restriction enzymes: B, *Bam*HI; Bc, *Bcl*I; Bg, *Bgl*II; E, *Eco*RI; K, *Kpn*I; M, *Mlu*I; P, *Pst*I; S, *Sac*I; Sa, *Sal*I.)

Figure 5B represents the organisation of the right part of the cluster of the biosynthesis genes of erythromycin in *Sac. erythraea* the ORFs 13 to 21 of which are indicated by arrows as well as a restriction map of the plasmids pBK6-12, pCN9, pNCO28, pNB49, pNCO62, pPSP4, pNCO62X and pBAB18. (Abbreviations of the restriction enzymes: B, *Bam*HI; Ba, *Bal*I; Bc, *Bcl*I; C, *Cla*I; E, *Eco*RI; K, *Kpn*I; N, *Nco*I; Ns, *Nsi*I; P, *Pst*I; Pv, *Pvu*II; S, *Sac*I; Sc, *Sca*I; Sh, *Sph*I; Sp, *Spe*I; X, *Xba*I; Xh, *Xho*I).

Figure 6A represents the construction diagram for the plasmid pBIIΔ.

Figure 6B represents a restriction map of the plasmid pUWL218.

Figure 6C represents a restriction map of the plasmid pBIIΔ.

Figure 7A represents the construction diagram of the

plasmid pdel88.

Figure 7B represents the construction diagram of the plasmid pdel88A.

Figure 7C represents the construction diagram of the plasmid pOBB.

Figure 7D represents the construction diagram and a restriction map of the plasmid pCIIIΔ.

Figure 8A represents the construction diagram of the plasmid pCIIΔ.

Figure 8B represents a restriction map of the plasmid pORT1.

Figure 8C represents a restriction map of the plasmid pCIIΔ.

Figure 9A represents the construction diagram of the plasmid pBIVΔ.

Figure 9B represents a restriction map of the plasmid pBIVΔ.

Figure 10A represents the construction diagram of the plasmid pBVΔ.

Figure 10B represents a restriction map of the plasmid pBVΔ.

Figure 11A represents the construction diagram of the plasmid pPSTI.

Figure 11B represents a restriction map of the plasmid pPSTI.

Figure 12A represents the construction diagram of the plasmid pXhoI.

Figure 12B represents a restriction map of the plasmid pXhoI.

Figure 13A represents the construction diagram of the plasmid pCIVΔ.

Figure 13B represents a restriction map of the plasmid pCIVΔ.

Figure 14A represents the construction diagram of the plasmid pCVΔ.

Figure 14B represents a restriction map of the plasmid pCVΔ.

Figure 15 represents the analysis by Southern blot of

the mutant strains BII92, CIII68, CII62, BIV87, BV88, CIV89 and CV90, compared with the "red variant" wild strain marked Wt. For each mutant, the restriction enzyme used is indicated below each blot and the size of the bands detected before each blot is estimated relative to the molecular weight markers λ-*Hin*dIII and λ-*Bst*EII (not detectable by auto-radiography).

Figure 16 represents the analysis by PCR of the mutant strains BII91, CIII68, CII62, BIV87, BV88, CIV89 and CV90, compared with the "red variant" wild strain marked Wt and with the plasmids pBIIΔ, pCIIIΔ, pCIIΔ, pBIVΔ, pBVΔ, PCIVΔ and pCVΔ used respectively in order to obtain the mutant by homologous recombination. The sizes of the bands detected by ethidium bromide staining are estimated relative to the molecular weight markers ΦX174-*Hae*III or λ-*Bst*EII. Figure 17 represents the analysis by TLC of the metabolites produced by the mutant strains BII92, CIII68, CII62, BIV87, BV88, CIV89 and CV90, compared with the standard products erythromycin A (Er A), erythronolide B (EB) and 3-α-mycarosyl erythronolide B (MEB).

Figure 18 represents the analysis by SDS-PAGE of the purification of the EryCIII protein successively after extraction with urea 7M (line 2), Q Sepharose chromatography (line 3), Superdex chromatography (line 4), Q source chromatography (line 6) with standard molecular weight markers (lines 1 and 5);

Figure 19 represents the analysis by TLC of the biological activity test of the EryCIII protein, by incubation with d-TDP-D-desosamine (line 2) or with d-TDP-D-desosamine and 3-α-mycarosyl erythronolide B (MEB) (line 3) compared with the MEB control (line 1) and with the erythromycin A control (line 4). The dotted lines mark the zones showing an antibiotic activity using an autobiogram on *B. pumilus*.

Figure 20 represents the localization of the six cosmids (cosAB35, cosAB76, cosAB87, cosAB67, cosAB63 and cosAB61) covering all of the cluster of the oleandomycin biosynthesis genes. The restriction fragments *Bam*HI (marked B)

hybridizing with the probes marked *str* M, D, E and the *Bam*HI fragments (3.5 kb and 2.7 kb) hybridizing with the probe *ery*CIII are shown.

Figure 21 represents the organisation of the right part of the cluster of the oleandromycin biosynthesis genes in *S. antibioticus* the different ORFs of which (marked *ole*P1, *ole*G1, *ole*G2, *ole*M, *ole*Y, *ole*P and *ole*B) are indicated by arrows as well as a restriction map of the plasmid pCO35-S and the insert of the plasmid pCO3 generated from pCO35-S. The double arrow indicates the insert corresponding to the sequence of Figure 22 (abbreviations of the restriction enzymes: B, *Bam*HI; Bg, *Bgl*II; K, *Kpn*I; S, *Sac*I; Sh, *Sph*I; the star indicates that it is not a unique site).

Figure 22 represents the nucleotide sequence (sequence of SEQ ID No.15) of the region covering the *ole*P1, *ole*G1, *ole*G2, *ole*M and *ole*Y oleandomycin biosynthesis genes and their deduced protein sequences.

EXAMPLE 1: cloning and sequencing the *ery*G-*ery*AIII region of the cluster of biosynthesis genes of erythromycin.

A genomic DNA fragment of *Sac. erythraea* NRRL 2338 having > 20 kb downstream from the *erm*E gene covering in particular ORFs 3 to 9 and corresponding to the clone λSE5.5 as well as the nucleotide sequence of a 4.5 kb fragment corresponding to the region of the *ery* cluster comprised between 3.7 kb and 8.0 kb from the 3' end of the *erm*E gene and comprising ORFs 3, 4, 5 and 6 were described by Haydock et al. (1991).

Taking account of the restriction map shown by Haydock et al. (1991), sub-clones were derived from clone λSE5.5 by sub-cloning restriction fragments in pUC19. The plasmids pKB22, pBK44, pBIISB and p*Eco*2 were thus generated according to Figure 5A in the following fashion:

From the DNA of clone λSE5.5 digested by the *Kpn*I restriction enzyme, plasmids pK62 and pK66 were directly constructed by sub-cloning of the *Kpn*I fragment of 5.8 kb in pUC19, the plasmid pK66 corresponding to the same *Kpn*I fragment sub-cloned with an reversed orientation of the insert relative to the vector. The plasmid pKB22 containing

an insert of 2.9 kb was then derived from plasmid pK66 by excision of the *Bam*HI-*Bgl*II fragment (2.9 kb) covering ORF8 as well as a part of ORFs 7 and 9 by digestion with the *Bam*HI and *Bgl*II restriction enzymes. In the same way, the plasmid pKB44 containing a insert of 2.9 kb was obtained from the plasmid pK62 by excision of the *Bam*HI-*Bgl*II fragment (2.9 kb) covering the *ery*G gene corresponding to ORFs 5 and 6 and the *ery*F gene corresponding to ORF4.

The plasmid pBIISB was derived from the plasmid pBK44 by sub-cloning in pUC19 of the *Sal*I fragment of 600 bp obtained from the plasmid pBK44 digested by the *Sal*I restriction enzyme (Figure 5A).

From the DNA of the clone λSE5.5 digested by the *Eco*RI restriction enzyme, the plasmid pEco2 was directly constructed by sub-cloning of the *Eco*RI fragment (2.2 kb) in pUC19.

The sub-clones pKB22, pBK44, pBIISB and pEco2 thus obtained were then sequenced. Analysis was carried out on the plasmidic DNA samples, purified beforehand on a Quiagen 100 column (Quiagen), on the ABI prism 377 automatic sequencer. The sequencing reactions were carried out by the Sanger method (1977) using conventional M13 primers or synthetic primers and fluorescent triphosphate didesoxynucleosides and Taq FS polymerase (Perkin Elmer) in the presence of 5 % of dimethylsulphoxide, the synthetic primers used having the following sequences:

C3R2	TCCTCGATGGAGACCTGCC	(SEQ ID No.22)
B2R1	GAGACCATGCCCAGGGAGT	(SEQ ID No.23)
C3S2	TCTGGGAGCCGCTCACCTT	(SEQ ID No.24)
C2R1	GACGAGGCCGAAGAGGTGG	(SEQ ID No.25)
C2S	GCACACCGGAATGGATGCG	(SEQ ID No.26)
fullC3S	CCGTCGAGCTCTGAGGTAA	(SEQ ID No.27)
fullC3R	GCCCGAGCCGCACGTGCGT	(SEQ ID No.28) and
C4	TGCACGCGCTGCTGCCGACC	(SEQ ID No.29).

Assembling of the sequence data was carried out with the Autoassembler™ software package (Applied Biosystem). The sequences were analyzed using the GCG set of software (Devereux 1984).

The nucleotide sequences obtained allowed the nucleotide sequence of 3412 bp of Figure 2 to be established (direct and complementary sequence of SEQ ID No.1) in which three ORFs (7, 8 and 9) were identified respectively from nucleotide 8957 to nucleotide 7959, from nucleotide 10219 to nucleotide 8957 and from nucleotide 11315 to nucleotide 10233 (numbered in Figure 2 from the *Bam*HI site situated at the 5' end of the *erm*E gene) (respectively complementary sequence of SEQ ID No.1 from nucleotide 48 to nucleotide 1046, from nucleotide 1046 to nucleotide 2308 and from nucleotide 2322 to nucleotide 3404) and corresponding respectively to the *ery*BII, *ery*CIII and *ery*CII genes according to Liu and Thorson (1994) the functional characterizations of which had not yet been identified. The three ORFs 7, 8 and 9 have the same orientation, reading being carried out from the 3'region of the *ery*AIII gene.

Specimens of *E. coli* XL1-blue containing the coding region for the above ORFs were deposited at the Collection Nationale de Cultures de Microorganismes (CNCM) INSTITUT PASTEUR, 25, Rue du Docteur Roux 75724 PARIS CEDEX 15 FRANCE, on the 16 July 1997:

- the plasmid pK62 comprising the coding sequence for ORF7, ORF8 and part of ORF9 under the number I-1897,
- the plasmid p***Eco***2 comprising the coding sequence for ORF9 and part of ORF8 under the number I-1899.

EXAMPLE 2: construction of the plasmid pBIIΔ.

An integration plasmid, called **pBIIΔ** and carrying a deletion in the *ery*BII gene coding for ORF7, was constructed according to the diagram in Figure 6A.

The *Bcl*I-*Bam*HI fragment of 598 bp was deleted in the plasmid pK62 obtained in Example 1 by digestion with the *Bcl*I and *Bam*HI enzymes. The resulting plasmid pBCK1 was then digested with the *Mlu*I and *Bgl*II restriction enzymes so as to delete a 853 bp fragment inside ORF7 from nucleotide 8011 to nucleotide 8863 of the sequence of Figure 2. After filling in the ends using the Klenow fragment of DNA polymerase I, the plasmid containing the deletion was religated and transformed in *E. coli* XL1-blue. From the plasmid p19BIIΔ

thus generated, the *Kpn*I-*Hin*dIII fragment (4.3 kb) which carries the deletion was sub-cloned in the plasmid pUWL218 (Figure 6B). The presence of the 853 bp deletion from nucleotide 8011 to nucleotide 8863 in the plasmid pBIIΔ thus generated (Figure 6C) was confirmed by sequencing.
The plasmid pBIIΔ was then transferred into the *E. coli* DH5αMRC strain, then used to transform *Sac. erythraea*.

EXAMPLE 3: construction of a *Sac. erythraea ery* BIIΔ strain (BII92).

The construction of a *Sac. erythraea* strain in which the *ery*BII gene carried an internal deletion such as that introduced into the plasmid pBIIΔ prepared in Example 2 and the integration process were carried out in the following fashion:

The preparation of the protoplasts was carried out according to the method described by Weber and Losick (1988), using PEG 3350 (Sigma) instead of PEG 1000 and a modified P buffer (called PT) containing $MgCl_2.6H_2O$ 28 mM and without PO_4H_2K instead of the P, L or T buffers described, according to the following operating conditions:

The cells (at least 10^8 spores) of *Sac. erythraea* "red variant" (a sample of which was deposited at the Collection Nationale de Cultures de Microorganismes (CNCM) INSTITUT PASTEUR, 25, Rue du Docteur Roux 75724 PARIS CEDEX 15 FRANCE, on the 16 July 1997 under the number I-1902) were grown in 50 ml of TBS medium for 3 to 5 days at 30°C, then washed with 10.3 % sucrose. The cells were resuspended in 50 ml of PT buffer containing 2 to 5 mg/ml of lysozyme (Sigma), then incubated at 30°C for 1 to 2 hours, with desegregation the mycelium mass every 15 minutes until transformation of at least 50 % of the mycelium to protoplasts. The protoplasts were washed with 50 ml of PT buffer, resuspended in 12.5 to 25 ml of the same buffer, slowly frozen then stored at -80°C by 200 µl aliquots.

For the transformation, one aliquot was thawed and 50 µl were removed then transferred into a 15 ml tube. One to 10 µg of plasmidic DNA pBIIΔ, prepared in Example 2 from the *E. coli* DH5αMRC strain were dissolved in 5 to 10 µl of TE buffer

(Tris HCl 10 mM pH 7.5, EDTA 1 mM) then deposited on the wall of the inclined tube to which was then added 0.5 ml of a solution of PEG 3350 in PT buffer prepared extemporaneously from an aqueous solution at 50 % which is diluted to one half in the 2 x PT buffer. After dilution with 3 to 5 ml of PT buffer then centrifugation at 2500 rpm for 15 mn, the pellet was dissociated in 0.5 ml of PT buffer and the suspension of transformed protoplasts thus obtained is immediately distributed onto 2 or 3 very dry R2T2 dishes(3 hours under a laminar flow hood). The dishes were then incubated at 32°C for 16 to 24 hours until the regeneration film of the protoplasts appears. From a stock solution of thiostrepton (Sigma) at 50 mg/ml in DMSO, an appropriate quantity was diluted in 0.5 to 1 ml of water then spread on the dishes so as to obtain a final concentration of 20 µg of thiostrepton/ml of gelose. After complete absorption of the antibiotic, the dishes were incubated at 32°C for 3 to 4 days, which allows visualisation of the transformants. The dishes were incubated for several more days until the spores were completely developed.

Selection of the integrants corresponding to the first recombination event (Figure 4) was carried out by replication of the sporulated dishes using velvet or by spreading a suspension of the spores onto R2T2 dishes containing thiostrepton then incubation at 32°C, which allows clones of potential integrants to grow.

For the selection of clones which have undergone a second recombination event (Figure 4), 5 to 10 thiostrepton-resistant clones obtained above were cultured in 8 ml of TSB liquid medium at 30°C for 3 to 4 days. 50 to 100 µl were removed and recultured under the same conditions. After 4 successive cycles of dilution and culture intended to encourage the loss of the thiostrepton resistance marker, protoplasts were prepared from the cells as indicated above, so as to expel the plasmid. The protoplasts were then spread onto R2T2 dishes so as to obtain individualized colonies the sensitivity to thiostrepton of which was determined by replication on R2T2 dishes containing thiostrepton.

Depending on the position of the second recombination event relative to the deletion site (Figure 4), it can be expected that the phenotype of the colonies which are sensitive to thiostrepton are of wild type or of mutated type carrying the deletion.

Among the colonies which are sensitive to thiostrepton, selection of the mutants having the *ery*⁻ phenotype was carried out by record of bacterial sensitivity to antibiotics on the *B. pumilus* ATCC 14884 strain which is sensitive to erythromycin. The *B. pumilus* strain was used as and an indicator strain for evaluating the production of erythromycin in biological tests by record of bacterial sensitivity to antibiotics. The colonies were spread using a platinum loop onto R2T2 dishes, then incubated for 3 to 4 days at 30°C. The agar areas where the mutant has grown at confluency are then removed with a punch then placed on A-Merck dishes covered with a 4 ml overlayer of 0.5 x A Merck (Antibiotic agar No.1 Merck) inoculated with a suspension of *B. pumilus* spores, then incubated overnight at 37°C.

The presence of the expected deletion in the chromosome of the mutant (deletion of 853 bp from nucleotide 8011 to nucleotide 8863 in Figure 2) was then confirmed by genomic analysis using Southern blot as well as by PCR in the following fashion:

For the analysis by Southern blot, the transfer of genomic DNA, digested beforehand with the appropriate restriction enzyme, on GeenscreenPlus (Dupont NEN) membranes was carried out in 0.4 M NaOH according to Ausubel et al. (1995). The hybridizations were carried out using as a probe the oligonucleotide labelled at its 5' end using [γ^{32}P]ATP (Amersham) and the polynucleotide kinase (Boehringer Mannheim) according to Sambrook et al. (1989), having the following sequence:

B2-S TTGGCGAAGTCGACCAGGTC (SEQ ID No.30)

corresponding to the DNA region at the start of the *eryG* gene situated from position 4118 to position 4137 of the sequence deposited in the EMBL base under the reference X60379 and described by Haydock et al. (1991). The hybridizations were

carried out with a rapid hybridization buffer (Amersham) and the following washing conditions: 2 x 5 mn, 2 x SSC, 20°C; 30 mn, 2 x SSC, 65°C; 30 mn, 0.1 x SSC, 20°C.

By Southern hybridization on the genomic DNA isolated according to Hopwood et al. (1985) then digested by the *Kpn*I restriction enzyme, a band of 5.8 kb from the "red variant" wild strain and a band of 4.9 kb from the mutant BII92 were detected. The results shown in Figure 15 indicate the presence in the mutant of a deletion of approximately 900 bp in this region of the chromosome.

For the analysis by PCR ,a 100 µl sample of a 3-day culture in TSB medium was centrifuged. The pellet obtained was resuspended in 10 µl of TSB medium, then used for amplification in the genAmp PCR system 9600 apparatus (Perkin Elmer Cetus). After heating the sample for 3 minutes at 94°C, the following amplification conditions were used: 94°C, 1 mn; 55°C, 1 mn; 72°C, 3 mn; 30 cycles; Ampli Taq polymerase (Perkin Elmer) in the presence of 10 % dimethylsulphoxide (v/v) followed by an extension of 3 mn at 72°C . The amplification was carried out using the oligonucleotide B2S above and the oligonucleotide having the following sequence:

B2-R GCCGCTCGGCACGGTGAACTTCA (SEQ ID No.31)

corresponding to the sequence of the complementary strand of the DNA region situated from position 8873 to position 8892 of the sequence of Figure 2 to which three nucleotides were added at the 5' end and which allows framing by PCR amplification of the region carrying the internal deletion at ORF7.

The analysis by PCR amplification on whole cells allowed a band of approximately 1 kb in the wild strain and a band of 0.16 kb in the mutant BII92 to be detected in an identical fashion to the signal obtained with the plasmid pBIIΔ. The results shown in Figure 16 confirm that the deletion of approximately 900 bp detected by the Southern analysis is identical to that carried by the plasmid pBIIΔ (853 bp).

The recombinant strain thus obtained, designated BII92, was then cultivated in order to identify the metabolites produced by the strain.

EXAMPLE 4: fermentation of the BII92 strain and identification of the secondary metabolites produced.

Extracts of the broth culture medium of the strain were analyzed by thin-layer chromatography (TLC) with erythromycin A, erythronolide B and 3-α-mycarosyl erythronolide B as standards.

The BII92 strain was cultured in a 50 ml erlenmeyer flask under conditions allowing an optimal production of erythromycin A and its derivatives which consist of carrying out a cell preculture at 28°C for 48 hours in EP1 medium (Solulys L-Corn steep liquor (Roquette frères) 5 g/l; defatted soya bean flour (Cargill) 10 g/l; CO_3Ca 2 g/l; NaCl 5 g/l; pH = 6.8; glucose sqf 15 g/l added after autoclaving), then a culture for 72 hours after dilution to 7 % v/v with EP2medium (defatted soya bean flour 10 g/l; CO_3Ca 0.2 g/l; Cl_2Co-6H_2O 1 mg/l; pH = 6,8-7.0; glucose sqf 20 g/l added after autoclaving).

The culture supernatant was then extracted at pH 9-10 with ethyl acetate. The organic phases were dried over $MgSO_4$, brought to dryness under reduced pressure then analyzed by TLC on 60 F254 silica gel (Merck) [dichloromethane/methanol (90:10, v/v) or isopropyl ether /methanol/NH_4OH at 25 % (75:35:2, v/v)]. Alternatively, the analysis was carried out by TLC on bonded silica gel plates of NH_2 F254 type (Merk) [butyl chloride/methanol (90:10, v/v)].

The chemical revelation of the plates was carried out by pulverization of a solution of p-anisaldehyde- 98 %sulphuric acid -ethanol (1:1:9, v/v), followed by heating for a few minutes at 80°C. The potential antibiotic activities were analyzed by direct bioautography of the TLC plates on agar seeded with B. *pumilus* ATCC 14884.

The results obtained by chemical revelation (Figure 17) shows that the BII92 strain preferentially accumulates erythronolide B as expected of an *ery*B mutant.

Minor metabolites of low mobility showing an antibiotic activity were also detected. These metabolites were extracted with ethyl acetate and identified by reversed-phase high-performance liquid chromatography (RP-HPLC) coupled with

mass spectrometry. The RP-HPLC was carried out on a column (250 x 4.6mm) of Kromasil C18 5µ using an acetonitrile/methanol/ammonium acetate 0.065 M pH 6.7 mixture (350:150:500, v/v) as the mobile phase, on a Waters chromatograph equipped with a Finningan TSQ 7000 mass spectrometer.

Alongside traces of erythromycin A, B, C and D, 4 minor metabolites called M1 to M4 were detected:

- M1 produces a parent peak at m/z 704 and fragmentation products at m/z 576 and m/z 158. The presence of desaminyl-erythronolide A (m/z 576) indicates that the difference in m/z of 30 compared with erythromycin A (m/z 734) or of 16 compared with erythromycin C (m/z 720) is carried by the neutral sugar residue. The proposed structure for M1 is 3"-C desmethyl-2",3"-ene-erythromycin C.
- M2 produces a parent peak at m/z 706 and fragmentation products at m/z 576 and m/z 158. The presence of desaminyl-erythronolide A (m/z 576) indicates that the difference in m/z of 28 compared to erythromycin A (m/z 734) or of 14 compared to erythromycin C (m/z 720) is carried by the neutral sugar residue. The proposed structure for M2 is 3"-C desmethyl-erythromycin C.
- M3 produces a parent peak at to m/z 690 and fragmentation products at m/z 560 and m/z 158. The presence of desaminyl-erythronolide B (m/z 560) indicates that the difference in m/z of 28 compared to erythromycin B (m/z 718) or of 14 compared to erythromycin D (m/z 704) is carried by the neutral sugar residue. The proposed structure for M3 is 3"-C desmethyl-erythromycin D.
- M4 produces a parent peak at m/z 720 and fragmentation products at m/z 576 and m/z 158. The profile is identical to that of erythromycin C (m/z 720) with the presence of desosaminylerythronolide A (m/z 576) and the loss of the aminated sugar residue (m/z 158), but the metabolite M4 does not have the same retention time in RP-HPLC as erythromycin. The proposed structure for M4 is 3"-C desmethyl-erythromycin A.

The detection by MS-MS of the minor metabolite M1 having

an unsaturated neutral sugar (3"-C desmethyl-2",3"-ene-erythromycin C) indicates that the *ery*BII gene codes for dTDP-4-keto-L-6-desoxyhexose 2,3-reductase in the biosynthesis route of dTDP-mycarose.

The BII92 strain was deposited at the Collection Nationale de Cultures de Microorganismes (CNCM) INSTITUT PASTEUR, 25, Rue du Docteur Roux 75724 PARIS CEDEX 15 FRANCE, on the 16 July 1997 under the number I-1903.

EXAMPLE 5: construction of the plasmid pCIIIΔ.

ORF8 may be translationally coupled to ORF7 situated downstream, an in phase deletion was introduced so as to avoid a polar effect. An integration plasmid, called pCIIIΔ which carries such a deletion, was constructed according to the diagram in Figure 7(A-D).

An *Sal*I deletion of 663 bp was introduced into ORF8 from nucleotide 9384 to nucleotide 10046 of the sequence of Figure 2 by sub-cloning in the plasmid pUC19 the two *Sal*I fragments (a: 794 bp and b: 631 bp shown in Figure 5A) isolated from the plasmid pBK44 obtained in Example 1 in order to generate the plasmid pdel88 (Figure 7A). The presence of the deletion of 663 bp was confirmed by sequencing. The plasmid pdel88 was then subjected to two additional sub-clonings so as to extend the chromosomal regions which can be used for the homologous recombination of the two sides of the deletion site. The *Sac*I fragment (450 bp) of the plasmid pdel88 was first replaced by the *Sac*I fragment (1.1 kb) of the plasmid pEco2 obtained in Example 1 in order to generate the plasmid pdel88A (Figure 7B). Then the *Eco*RI fragment (1.5 kb) carrying the deletion in ORF8 was isolated from the plasmid pdel88A and used to replace the *Eco*RI fragment (1.66 kb) carrying the intact ORF in the plasmid pOBB. The plasmid pOBB, represented in Figure 7C, corresponds to the plasmid pBK44 prepared in Example 1 in the site *Pst*I of which there was sub-cloned the fragment *Pst*I of 4 kb of the plasmid pIJ486 obtained by digestion by the *Pst*I restriction enzyme and carrying of the streptomyces replication origin as well as the thiostrepton resistance gene. The resultant plasmid pCIIIΔ carries chromosomal regions for the homologous

recombination of 1.27 kb and 1.38 kb upstream and downstream respectively of the deletion site. The plasmid pCIIIΔ thus obtained (Figure 7D) was then transferred into the *E. coli* DH5αMRC strain, then used for transforming *Sac. erythraea*.

EXAMPLE 6: construction of a *Sac. erythraea* eryCIIIΔ strain (CIII68).

A strain in which the *ery*CIII gene carries an internal deletion such as that introduced into the plasmid pCIIIΔ obtained in Example 5 was prepared by transformation of the protoplasts of *Sac. erythraea* with the plasmid pCIIIΔ.

The preparation of the protoplasts, the integration process and the selection of the mutants having the *ery*$^-$ phenotype were carried out as in Example 3.

In addition, the presence of the expected deletion in the chromosome (deletion of 663 bp from nucleotide 9384 to nucleotide 10046 of the sequence of Figure 2) was confirmed by genomic analysis by Southern blot as well as by PCR amplification according to the conditions described in Example 3.

Using Southern hybridization on the genomic DNA digested by the *Eco*RI restriction enzyme, using as probe the oligonucleotide having the following sequence

C3-S ATGCGCGTCGTCTTCTCCTCCATG (SEQ ID No.32)

corresponding to the complementary strand of the DNA region situated from position 10196 to position 10219 of the sequence in Figure 2, a band of 2.2 kb from the wild strain and a band of 1.5 kb from the mutant CIII68 were detected. The results shown in Figure 15 indicate the presence in the mutant of a deletion of approximately 700 bp in this region of the chromosome.

The PCR amplification was carried out using the above oligonucleotide C3-S and the oligonucleotide having the following sequence

C3-R TCATCGTGGTTCTCTCCTTCC (SEQ ID No.33)

corresponding to the sequence situated from position 8954 to position 8974 of the sequence in Figure 2 allowing framing by PCR amplification of the region carrying the internal deletion at ORF8. Analysis by PCR amplification allowed a

band of approximately 1.2 kb in the wild strain and a band of approximately 0.6 kb in the mutant CIII68 to be detected in an identical fashion to the signal obtained with pBIIΔ. The results shown in Figure 16 confirm that the deletion of approximately 700 bp detected by the Southern analysis is identical to that carried by the plasmid pCIIIΔ (633 bp).

The recombinant strain thus obtained, designated CIII68, was then cultivated in order to identify the metabolites produced by the strain.

<u>EXAMPLE 7</u>: fermentation of the CIII68 strain and identification of the secondary metabolites produced.

The culture of the CIII68 strain and the analyses by TLC followed by bioautography were carried out according to the conditions indicated in Example 4.

The TLC results (Figure 17) show that the CII68 strain preferentially accumulates 3-α-mycarosyl erythronolide B as well as small quantities of erythronolide B as expected of an *ery*C mutant.

The *ery*CIII sequence has a strong homology with other putative glycosyltransferases such as DauH (43 % identity at the protein level) and DnrS (47 % identity) involved in the biosynthesis of daunorubicin in *S. peucetius* (Otten et al., 1995) and in *Streptomyces* sp C5 (Dickens et al., 1996) as well as TylM2 (50 % identity) involved in the transfer of mycaminose onto tylactone in the biosynthesis route of tylosine in *S. fradiae* (Gandecha et al., 1997).

These observations indicate that *ery*CIII gene codes for desosaminyltransferase in the biosynthesis route of erythromycin.

<u>EXAMPLE 8</u>: construction of the plasmid pCIIΔ .

An integration plasmid, called pCIIΔ and carrying a deletion in the *ery*BII gene coding for ORF9, was constructed according to the diagram in Figure 8A.

The plasmid pK23 (Figure 5A) was obtained by sub-cloning in pUC19 of the *Kpn*I fragment of 10 kb isolated from the DNA of the clone λSE5.5 digested by the *Kpn*I restriction enzyme.

Firstly, the shuttle vector pORT1, shown in Figure 8B, was obtained by sub-cloning the *Pst*I fragment of 4kb isolated

by digestion of the plasmid pIJ486 with the *Pst*I restriction enzyme including the thiostrepton resistance gene and the Streptomyces replicon, in the *Pst*I site of pUC19.

An out-of-phase deletion of 304 bp was introduced into ORF9 from nucleotide 10881 to nucleotide 11184 of the sequence of Figure 2 by sub-cloning the ***Sac*I-*Kpn*I** fragment (1.1 kb) of the plasmid pK23 with the *Eco*RI-*Kpn*I fragment (1.7 kb) of the plasmid p*Eco*2 obtained in Example 1 in the above plasmid pORT1 digested beforehand with the *Sac*I and *Eco*RI restriction enzymes. The integration plasmid pCII thus obtained (Figure 8C) was then transferred into the *E. coli* DH5αMRC strain, then used to transform *Sac. erythraea*.

__EXAMPLE 9__: construction of a *Sac. erythraea ery*CIIΔ strain (CII62).

A strain in which the *ery*CII gene carries an internal deletion such as that introduced into the plasmid pCIIΔ obtained in Example 8 was prepared by transformation of the protoplasts of *Sac. erythraea* with the plasmid pCIIΔ.

The preparation of the protoplasts, the integration process and the selection of the mutants having the ery^- phenotype were carried out as in Example 3.

In addition, the presence of the expected deletion in the chromosome (deletion of 304 bp from nucleotide 10881 to nucleotide 11184 of the sequence of Figure 2) was confirmed by genomic analysis by Southern blot as well as by PCR according to the conditions described in Example 3.

Using Southern hybridization on the genomic DNA digested by the *Eco*RI restriction enzyme, using as probe the oligonucleotide C3-S having the above sequence, a band of 2.2 kb from the wild strain and a band of 1.8 kb from the mutant CII62 were detected. The results shown in Figure 15 indicate the presence in the mutant of a deletion of approximately 400 bp in this region of the chromosome.

The PCR amplification was carried out using the oligonucleotide having the following sequence

C2-S GGAATTCATGACCACGACCGATC (SEQ ID No.34)

corresponding to the complementary strand of the DNA region of the end of the *ery*AIII gene situated from position 20258

to position 20280 of the sequence deposited in the EMBL base under the reference X62569 and described by Bevitt et al., 1992 and the oligonucleotide having the following sequence

C2-R CGCTCCAGGTGCAATGCCGGGTGCAGGC (SEQ ID No.35)

corresponding to the sequence situated from position 10558 to position 10585 of the sequence of Figure 2 allowing framing by PCR amplification of the region carrying the internal deletion at ORF9. Analysis by PCR amplification allowed a band of approximately 760 bp in the wild strain and a band of approximately 460 bp in the mutant CII62 to be detected in an identical fashion to the signal obtained with the plasmid pCIIΔ. The results shown in Figure 16 confirm that the deletion of approximately 400 bp detected by the Southern analysis is identical to that carried by the plasmid pCIIΔ (304 bp).

The recombinant strain thus obtained, designated CII62, was then cultivated in order to identify the metabolites produced by the strain.

EXAMPLE 10: fermentation of the CII62 strain and identification of the secondary metabolites produced.

The culture of the CII62 strain and the analyses by TLC followed by bioautography were carried out according to the conditions indicated in Example 4.

The TLC results (Figure 17) show that the CII62 strain preferentially accumulates 3-α-mycarosyl erythronolide B as well as small quantities of erythronolide B as expected of an *ery*C mutant.

The *ery*CII sequence has a strong homology with genes involved in the biosynthesis routes for daunosamine (DnrQ, 38 % identity at the protein level, Otten et al., 1995) and mycaminose (protein coded by ORF1*, 40 % identity at the protein level, Gandecha et al., 1997) which also need to transfer a keto group in position 3 from an adjacent carbon.

These observations indicate that the *ery*CII gene codes for dTDP-4-keto-D-6-desoxyhexose 3,4-isomerase in the biosynthesis route of dTDP-desosamine.

EXAMPLE 11: cloning and sequencing of the *ery*AI-*ery*K region of the cluster of biosynthesis genes of erythromycin.

Cosmids containing the *ery*AI-*ery*K region of the cluster of ***ery*** genes such as cosmid Cos6B, were isolated by screening a genomic DNA library of *Sac. erythraea* in the cosmidic vector pWE15 (Stratagene) using as probe a DNA fragment of 13.2 kb comprising the totality of the *ery*AI gene and corresponding to the DNA region comprised between the *Nco*I site situated at position 44382 of the sequence of Figure 3 and the *Nco*I site situated at position 392 of the X62569 sequence (Bevitt et al., 1992). The probe was prepared in the following fashion: Firstly, the *Nco*I fragment of 13.2 kb was isolated from the plasmid pBK25 described by Bevitt et al., 1992 and sub-cloned in the *Sma*I site of pUC18 after filling in the *Nco*I ends with the Klenow fragment. From the plasmid pNCO12 thus generated, the fragment of 13.2 kb was isolated by digestion with the *Nco*I restriction enzyme.

The cosmid cos6B thus obtained was digested by the *Nco*I restriction enzyme and the resulting fragments of 2.8 kb and 6.1 kb were cloned in the *Nco*I site of the vector Litmus28 generating respectively the plasmids pNCO28 and pNCO62 shown in Figure 5B.

The plasmid pNCO28 was sequenced by the generation of sub-clones using exonuclease III according to the instructions given by the supplier for the Erase-a-Base Kit (Promega) by digesting with the restriction enzymes *Sac*I/*Xba*I and *Nsi*I/*Bam*HI respectively for the reverse direction. The sequence was completed using as primers the synthetic oligonucleotides having the following sequences

644 GATCACGCTCTTCGAGCGGCAG (SEQ ID No.36)
645 GAACTCGGTGGAGTCGATGTC (SEQ ID No.37) and
650 GTTGTCGATCAAGACCCGCAC (SEQ ID No.38)

For the sequencing of plasmid pNCO62, templates were generated by sonication of the DNA according to Bankier et al. (1987) using as primers pUC18 such as vector. The sequence was completed using the synthetic oligonucleotides having the following sequences:

646 CATCGTCAAGGAGTTCGACGGT (SEQ ID No.39)
647 TGCGCAGGTCCATGTTCACCACGTT (SEQ ID No.40)
648 GCTACGCCCTGGAGAGCCTG (SEQ ID No.41)

649 GTCGCGGTCGGAGAGCACGAC (SEQ ID No.42) and
874 GCCAGCTCGGCGACGTCCATC (SEQ ID No.43).

The *Nco*I junctions were sequenced using as template the DNA of the cosmid cos6B obtained above the regions of which covering the *Nco*I sites were sequenced using the primers having the sequences 644 and 645 indicated above.

In addition, a ***ClaI-NcoI*** fragment of 0.9 kb, containing the start of the sequence of the *ery*AI gene and the 5' part of ORF13, was cloned in pUC18. This fragment was prepared in the following fashion: The plasmid pBK6-12 represented in Figure 5B was first generated by sub-cloning in the phagmide pTZ18R of the *Kpn*I fragment of 4.5 kb isolated from the plasmid pBK25 described by Bevitt et al., 1992. The sub-clone pCN9 was then generated by sub-cloning of the *Cla*I-*Nco*I fragment of 0.9 kb isolated from the plasmid pBK6-12 in the *Sma*I site of pUC19, after filling in the ends using the Klenow fragment. The plasmid pCN9 thus obtained (Figure 5B) was sequenced. Templates were generated by sonication of the DNA according to Bankier et al. (1987) using pUC18 as vector. The sequence was completed using as probe the oligonucleotide having the following sequence:

875 CGACGAGGTCGTGCATCAG (SEQ ID No.44).

The sequencing of the DNA is carried out by the Sanger method (1977) using an automated sequencer on the double strand DNA templates with the Applied Biosystem 373 A sequencer. The assembling of the sequence data was carried out with SAP software(Staden, 1984). The sequences were analyzed using GCG software (Devereux, 1984).

The nucleotide sequences obtained allowed the nucleotide sequence of 8160 bp of Figure 3 (sequence of SEQ ID No.6) to be established in which seven ORFs (13-19) were identified respectively from nucleotide 43841 to nucleotide 44806, from nucleotide 44809 to nucleotide 46053, from nucleotide 46109 to nucleotide 46819, from nucleotide 46907 to nucleotide 48436, from nucleotide 48436 to nucleotide 49638, from nucleotide 49679 to nucleotide 51145 and from nucleotide 51177 to nucleotide 51755 (numbered in Figure 3 from the *Bam*HI site situated at the 5' end of the *erm*E gene)

(respectively sequence of SEQ ID No.6 from nucleotide 242 to nucleotide 1207, from nucleotide 1210 to nucleotide 2454, from nucleotide 2510 to nucleotide 3220, from nucleotide 3308 to nucleotide 4837, from nucleotide 4837 to nucleotide 6039, from nucleotide 6080 to nucleotide 7546 and from nucleotide 7578 to nucleotide 8156) and corresponding respectively to the *ery*BIV, *ery*BV, *ery*CVI, *ery*BVI, *ery*CIV, *ery*CV and *ery*BVII genes, according to Liu and Thorson (1994) the functional characterizations of which had not yet been identified. The seven ORFs (13-19) are in the same direction, reading being carried out from the 5'region of the *ery*AI gene.

Specimens of *E. coli* XL1-blue containing the coding region of the above ORFs were deposited at the Collection Nationale de Cultures de Microorganismes (CNCM) INSTITUT PASTEUR, 25, Rue du Docteur Roux 75724 PARIS CEDEX 15 FRANCE, on the 16 July 1997:

- the plasmid pBK6-12 comprising the coding sequence for ORF13 and for part of ORF14 under the number I-1898
- the plasmid pNCO28 comprising the coding sequence for ORFs 14 and 15 as well as for part of ORFs 13 and 16 under the number I-1901 and
- the plasmid pNCO62 comprising the coding sequence for ORFs 17, 18 and 19 as well as for part of ORF16 under the number I-1900.

EXAMPLE 12: construction of the plasmid pBIVΔ.

ORF13 being translationally coupled to ORF14 situated downstream, an in-phase deletion had to be introduced. An integration plasmid, called pBIVΔ and carrying this deletion, was constructed according to the diagram of Figure 9A.

The plasmid pPSP4 (Figure 5B) was first constructed by sub-cloning of the *Pvu*II-*Spe*I fragment (2.7 kb) isolated from the plasmid pBK6-12 obtained in Example 11 and the *Spe*I-*Pst*I fragment (1.6 kb) isolated from the plasmid pNCO28 obtained in Example 11 in the vector pUC19 digested beforehand using the *Sma*I and *Pst*I restriction enzymes.

From the plasmid pPSP4, the plasmid p19BIVΔ was generated by deleting the *Bcl*I-*Nco*I fragment of 510 bp internal to ORF13 and by substituting for it 45 bp

originating from a synthetic adaptor of 54 bp. This adaptor was generated by pairing the 2 complementary oligonucleotides having the following sequences

SEQ A

AATTGATCAAGGTGAACACGGTCATGCGCAGGATCCTCGAGCGGAACTCCATGGGG (SEQ ID No.45) and

SEQ B

CCCCATGGAGTTCCGCTCGAGGATCCTGCGCATGACCGTGTTCACCTTGATCAATT (SEQ ID No.46)

creating a *Bcl*I site and an *Nco*I site framing the sequence of 45 bp.

For the pairing, the two oligonucleotides were brought to a 1.8 µM final concentration in the hybridization buffer NaCl 50 mM, Tris, HCl 20 mM pH 7.4, $MgCl_2.6H_2O$ 2 mM, heated for 5 mn at 100°C then slowly cooled down to ambient temperature. After digestion with the *Nco*I and *Bcl*I restriction enzymes, a ligature was carried out in the plasmid pPSP4 the *Bcl*I-*Nco*I fragment of 510 bp of which had been eliminated beforehand. From the plasmid p19BIVΔ thus generated, the *Sac*I-*Eco*RI fragment (2.2 kb) carrying the modified ORF13 was sub-cloned in the plasmid pUWL218 digested beforehand with the *Sac*I and *Eco*RI restriction enzymes. The integration plasmid pBIVΔ thus obtained (Figure 9B) was then transferred into the *E. coli* DH5αMRC strain, then used to transform *Sac. erythraea*.

EXAMPLE 13: construction of a *Sac. erythraea ery*BIVΔ strain (BIV87).

A strain in which the *ery*CIII gene carries an internal deletion such as that introduced into the plasmid pBIVΔ obtained in Example 12 was prepared by transformation of the protoplasts of *Sac. erythraea* with the plasmid pBIVΔ.

The preparation of the protoplasts, the integration process and the selection of the mutants having the *ery*⁻ phenotype were carried out as in Example 3.

In addition, the presence of the expected deletion in the chromosome (deletion of 510 bp from nucleotide 43872 to nucleotide 44382 of the sequence of Figure 3) and its replacement by the synthetic sequence of 45 bp was confirmed

by genomic analysis by Southern blot as well as by PCR according to the conditions described in Example 3.

Using Southern hybridization on the genomic DNA digested by the *XHo*I restriction enzyme, using as probe the B4-R oligonucleotide having the following sequence

B4-R AACTCGGTGGAGTCGATGTCGTCGCTGCGGAA (SEQ ID No.47)

corresponding to the complementary strand of the DNA region situated from position 44687 to position 44718 of the sequence of Figure 3, a band of 5.4 kb from the wild strain and a band of 2.7 kb from the mutant BIV87 were detected. The results shown in Figure 15 indicate the presence of an additional *Xho*I site at a distance of 2.7 kb upstream of the *Xho*I site situated at position 47114 of the sequence of Figure 3 thus confirming the incorporation of the adaptor in the mutant chromosome, as expected by the incorporation of the above synthetic adaptor used to generate the plasmid pBIVΔ.

The PCR amplification was carried out using the oligonucleotide having the following sequence

B4-S CAATATAGGAAGGATCAAGAGGTTGAC (SEQ ID No.48)

corresponding to the DNA region situated from position 43652 to position 43678 of the sequence of Figure 3 and the oligonucleotide B4-R having the sequence indicated above, allowing framing by PCR amplification of the region carrying the internal deletion at ORF13. Analysis by PCR amplification allowed a band of approximately 1 kb in the wild strain and a band of approximately 500 bp in the mutant BIV87 to be detected in an identical fashion to the signal obtained with the plasmid p BIV87 (Figure 16).

The whole of the Southern analysis and PCR results confirm the presence of the deletion of 510 bp and of the synthetic adaptor at the level of the chromosome of the mutant.

The recombinant strain thus obtained, designated BIV87, was then cultivated in order to identify the metabolites produced by the strain.

EXAMPLE 14: fermentation of the BIV87 strain and identification of the secondary metabolites produced.

The culture of the BIV87 strain and the analyses by TLC were carried out according to the conditions indicated in Example 4.

The TLC results (Figure 17) show that the BIV87 strain preferentially accumulates erythronolide B as expected of an *ery*B mutant.

Minor metabolites with low mobilities showing an antibiotic activity were also detected. These metabolites were extracted and analyzed by RP-HPLC coupled with mass spectrometry as described in Example 4.

The mass spectrum results show that modified forms of erythromycin A, B, C and D were produced. A major metabolite and 3 minor metabolites were detected.

The major metabolite M5 produces a parent peak at m/z 702 with dehydration and fragmentation products at m/z 684, m/z 560 and m/z 158 and corresponds to the elimination of 2 hydrogen atoms in erythromycin D (m/z 704, m/z 686). The presence of desosaminyl erythronolide B (fragment m/z at 560) indicates that the difference in mass is carried by the neutral sugar. The proposed structure for this metabolite is 4"-keto erythromycin D.

The minor metabolites also produce a profile with a difference of 2 in the m/z values respectively:

- M6 (m/z at 718, m/z 700, m/z 576, m/z 158) instead of m/z 720, m/z 702 for erythromycin C;
- M7 (m/z at 732, m/z 714, m/z 576, m/z 158) instead of m/z 734, m/z 716 for erythromycin A;
- M8 (m/z at 716, m/z 698, m/z 560, m/z 158) instead of m/z 718, m/z 700 for erythromycin B.

The proposed structures are respectively 4"-keto erythromycin C for M6, 4"-keto erythromycin A for M7 and 4"-keto erythromycin B for M8.

These observations indicate that the *ery*BIV gene codes for dTDP-4-keto-L-6-desoxyhexose 4-reductase in the biosynthesis route of dTDP-mycarose.

The BIV87 strain was deposited at the Collection Nationale de Cultures de Microorganismes (CNCM) INSTITUT PASTEUR, 25, Rue du Docteur Roux 75724 PARIS CEDEX 15 FRANCE,

on the 16 July 1997 under the number I-1904

EXAMPLE 15: construction of the plasmid pBVΔ.

An integration plasmid, called pBVΔ and carrying a deletion in the *ery*BV gene coding for ORF14, was constructed according to the diagram in Figure 10A.

A deletion of 726 bp was generated in ORF14 from nucleotide 44963 to nucleotide 45688 of the sequence of Figure 3 by ligation of the *Bcl*I-*Kpn*I fragment (1.1 kb) isolated from the plasmid pBK6-12, obtained in Example 11, with the *Kpn*I-*Bam*HI fragment (1.1 kb) isolated from the plasmid pNCO28, obtained in Example 11, in the plasmid pUWL218 digested beforehand by the *Bam*HI restriction enzyme. The integration plasmid pBVΔ thus obtained (Figure 10B) was then transferred into the *E. coli* DH5αMRC strain, then used to transform *Sac. erythraea*.

EXAMPLE 16: construction of a *Sac. erythraea ery*BVΔ strain (BV88).

A strain in which the *ery*BV gene carries an internal deletion such as that introduced into the plasmid pBVΔ obtained in Example 15 was prepared by transformation of the protoplasts of *Sac. erythraea* with the plasmid pBVΔ.

The preparation of the protoplasts, the integration process and the selection of the mutants having the *ery*$^-$ phenotype were carried out as in Example 3.

In addition, the presence of the expected deletion in the chromosome (deletion of 726 bp from nucleotide 44963 to nucleotide 45688 of the sequence of Figure 3) was confirmed by genomic analysis by Southern blot as well as by PCR amplification according to the conditions described in Example 3.

Using Southern hybridization on the genomic DNA digested by the *Nco*I restriction enzyme, using as probe the oligonucleotide having the following sequence:

B5-R TCCGGAGGTGTGCTGTCGGACGGACTTGTCGGTCGGAAA (SEQ ID No.49)

corresponding to the complementary strand of the DNA region situated from position 46060 to position 46098 of the sequence of Figure 3, a band of 2.7 kb from the wild strain and a band of 2.0 kb from the mutant BV88 were detected. The

results shown in Figure 15 indicate in the mutant the presence of a deletion of about 700 bp in this region of the chromosome.

The PCR amplification was carried out using the oligonucleotide having the following sequence:
B5-S AGGAGCACTAGTGCGGGTACTGCTGACGTCCTT (SEQ ID No.50)
corresponding to the DNA region situated from position 44799 to position 44831 of the sequence of Figure 3 and the oligonucleotide B5-R having the sequence indicated above, allowing framing by PCR amplification of the region carrying the internal deletion at ORF14. Analysis by PCR amplification has allowed a band of approximately 1.3 kb in the wild strain and a band of approximately 570 bp in the mutant BV88 to be detected in an identical fashion to the signal obtained with the plasmid pBV88 (Figure 16). These results confirm that the deletion of 710 bp detected by Southern analysis is identical to that carried by the plasmid pBVΔ (726 bp).

The recombinant strain thus obtained, designated BV88, was then cultivated in order to identify the metabolites produced by the strain.

EXAMPLE 17: fermentation of the BV88 strain and identification of the secondary metabolites produced.

The culture of the BV88 strain and the analyses by TLC were carried out according to the conditions indicated in Example 4.

The TLC results (Figure 17) show that the BV88 strain preferentially accumulates erythronolide B as expected of an *ery*B mutant.

Minor metabolites with low mobilities were also detected then extracted and identified by RP-HPLC coupled with mass spectrometry under the conditions used in Example 4.

The mass spectrum shows the presence of a metabolite having a parent peak at m/z 560 and dehydration and fragmentation products at m/z 542 and m/z 158 the proposed structure for which is desosaminyl erythronolide B.

The *ery*BV sequence shows a strong homology with other glycosyltransferases as well as with the *ery*CIII gene above (60.7 % identity at the nucleotide level, 44 % at the protein

level).

These observations indicate that the *ery*BV gene codes for the mycarosyltransferase involved in the biosynthesis of erythomycin.

EXAMPLE 18: construction of a plasmid pCVIΔ (pPSTI).

An integration plasmid, called pPSTI and carrying a deletion in the *ery*CVI gene coding for ORF15, was constructed according to the diagram in Figure 11A in the following fashion:

Firstly, the plasmid pNB49 was generated by treatment with the exonuclease III of the plasmid pNCO28 obtained in Example 11 digested beforehand by the *Nsi*I and *Bam*HI restriction enzymes. The plasmid pNB49 (Figure 5B) containing the nucleotides 44382 to 46562 of the sequence of Figure 3, was then digested with the help of the *Pst*I restriction enzyme then treated with Mung Bean nuclease (NE Biolabs) as described by Sambrook et al. (1989). After religation and transformation in *E. coli* XL1-Blue, the ampicillin-resistant colonies were selected by restriction analysis with the *Pst*I enzyme. The loss of the *Pst*I site was confirmed by sequencing a clone using the reverse M13 primer and deletion from nucleotide 46364 of the sequence of Figure 3 was observed creating a phase change in ORF15 in the plasmid pNB49ΔPst thus generated. The plasmid pIJ702 digested with the *Bgl*II restriction enzyme was then ligated with the *Bgl*II site of the plasmid pNB49ΔPst generating the plasmid pPSTI. The orientation of pIJ702 in pPSTI was confirmed by the presence of a DNA fragment having 0.9 kb after digestion with the *Sph*I restriction enzyme. The integration plasmid pPSTI (Figure 11B) thus obtained was transferred into the *E. coli* DH5αMRC strain, then used to transform *Sac. erythraea*.

EXAMPLE 19: construction of a *Sac. erythraea* *ery*CVIΔ strain (Pst10).

A strain in which the *ery*BV gene carries an internal deletion such as that introduced into the plasmid pPSTI obtained in Example 18 was prepared by transformation of the protoplasts of *Sac. erythraea* with the plasmid pPSTI.

The preparation of the protoplasts and the integration process were carried out as in Example 3.

The selection of the mutants having the *ery*⁻ phenotype was carried out as in Example 3 using a *B. subtilis* strain sensitive to erythromycin instead of a *B. pumilus* strain as indicator strain. The *B. subtilis* strain ATCC 6633 was used to evaluate the production of erythromycin in biological tests on agar dishes in M1-102 medium inoculated with the mutant to be analyzed and incubated for 3 days at 30°C. Agar zones covered with bacteria were then removed with the punch then placed on 2 x TY dishes covered with an overlayer of 5 ml of agar in TY medium containing 200 µl of a culture of *B. subtilis* ATCC 6633, then incubated overnight at 37°C.

The absence of production of erythromycin was also evaluated in the presence of added precursors such as erythronolide B or 3-α-mycarosyl erythronolide B by application of 10 µl of a 10 mM solution of each metabolite on the cut out agar zones followed by incubation at 30°C overnight before re-covering the dishes with the culture of *B. subtilis* as indicated above. The *Sac. erythraea* "red variant" wild strain was used as a control.

After transformation of the protoplasts with the plasmid pPSTI and selection of the thiostrepton-resistant colonies, integration in the chromosome was confirmed by Southern analysis according to the general methods described in Example 3.

A DNA fragment of 1269 bp corresponding to ORF14 generated by PCR using the synthetic oligonucleotides having the following sequences:

14-1 GGGGGATCCCATATGCGGGTACTGCTGACGTCCTTCG (SEQ ID No.51)
and
14-2 GAAAAGATCTGCCGGCGTGGCGGCGCGTGAGTTCCTC (SEQ ID No.52)
was used as probe.

The 14-1 oligonucleotide was designed to introduce a *Bam*HI site and a *Nde*I site upstream of the sequence corresponding to the DNA region situated from position 44811 to position 44833 of the sequence of Figure 3.

The 14-2 oligonucleotide was designed to introduce a

*Bgl*II site downstream of the sequence corresponding to the complementary strand of the DNA region situated from position 46027 to position 46053 of the sequence of Figure 3. The chromosomal DNA digested beforehand with the *Cla*I and *Pst*I restriction enzymes showed the expected bands of 4 kb and 7 kb from the integrant while the wild strain presented the expected band of 3 kb.

After repeated cultures of the integrants, the individualized colonies obtained were analyzed for sensitivity to thiostrepton and production of erythromycin, then the integration of the expected deletion (deletion from nucleotide 46364 of the sequence of Figure 3) into the chromosome of a mutant clone having the *ery*⁻ phenotype (Pst10) was confirmed by Southern analysis. The chromosomal DNA, isolated from the wild strain and the Pst10 mutant respectively, was digested with the *Pst*I restriction enzyme. Hybridization with the *Pst*I-*Nco*I probe of 0.8kb (nucleotides 46368 to 47142 of the sequence of Figure 3) produced the expected profile with a *Pst*I band of 1kb corresponding to nucleotides 46368 to 47397 of the sequence of Figure 3 from the wild strain and with a band of >20 kb from the mutant. The loss of the *Pst*I site at position 46368 above was also shown after double digestion by the *Pst*I and *Nco*I enzymes, resulting in a *Pst*I-*Nco*I band of 0.8 kb (nucleotide 46368 to 47142) from the wild strain and a ***Nco*I** band of 2.8 kb (nucleotide 44382 to 47142) with the mutant.

The recombinant strain thus obtained, designated Pst10, was then cultivated in order to identify the metabolites produced.

EXAMPLE 20: fermentation of the Pst10 strain and identification of the secondary metabolites produced.

The Pst10 strain was cultivated in the sucrose-succinate medium described by Caffrey et al. (1992) for 3 days at 30°C. The culture supernatant was then extracted at pH 9 with ethyl acetate. The organic phases obtained were dried over Mg_2SO_4 then brought to dryness under reduced pressure. The residue was dissolved in the acetonitrile-water mixture (1:1, v/v), then analyzed by mass spectrometry on a BioQ

(Micromass, Manchester, UK) or Finningan LCQ (Finningag, CA) spectrometer.

The production of erythomycin A (m/z 734 and m/z 716) was not observed but the presence of erythronolide B (MK+: m/z 441 and MNa+: m/z 425) as well as of 3-α-mycarosyl erythronolide B (MK+: m/z 585 and MNa+: m/z 569) demonstrated characterizes the Pst10 strain as an *ery*C mutant.

The *ery*CVI sequence has a strong homology with other methyltransferases such as SnoX involved in the biosynthesis of nogalamycin in *S. nogalater* (accession number EMBL S52403) (55.5 % identity at the protein level), TylM1 involved in the biosynthesis of tylosine in *S. fradiae* (accession number EMBL X81885) (65 % identity at the protein level) and SrmX involved in the biosynthesis of spiramycin in *S. ambofaciens* (accession number EMBL S25204) (52.8 % identity at the protein level).

These observations indicate that the *ery*CVI gene codes for the dTDP-D-6-desoxyhexose 3-N-methyltransferase involved in the biosynthesis route of dTDP-D-desosamine.

EXAMPLE 21: construction of a plasmid pBVIΔ (pXhoI).

An integration plasmid, called pXhoI and carrying a deletion in the *ery*BVI gene coding for ORF16, was constructed according to the diagram in Figure 12A in the following fashion:

The *Nco*I-*Xho*I fragment (3.1 kb) of the plasmid pNCO62 obtained in Example 11 and containing the nucleotides 47142 to 50254 of the sequence of Figure 3 was sub-cloned in the *Nco*I and *Xho*I sites of the plasmid Litmus 28. The plasmid pNCO62X (Figure 5B) thus generated was digested with the *Pst*I restriction enzyme then treated with DNA polymerase T4 (Boehringer Mannheim). After religation and transformation in *E. coli* XL1-Blue, the loss of the *Pst*I site at nucleotide 47397 of the sequence of Figure 3 was confirmed by sequencing and a deletion of 60 bp from nucleotide 47337 to nucleotide 47397 of the sequence of Figure 3 were observed. The plasmid pIJ702 digested with the *Bgl*II restriction enzyme was then ligated to the *Bgl*II site of this construction. The orientation of pIJ702 in the construction was confirmed by

the presence of a DNA fragment having 4.3 kb after digestion with the *Xho*I restriction enzyme. The integration plasmid **pXhoI** (Figure 12B) thus obtained was used to transform *Sac. erythraea*.

EXAMPLE 22: construction of a *Sac. erythraea* strain *ery*BVIΔ (Xho91).

A strain in which the *ery*BVI gene carries an internal deletion such as that introduced into the plasmid pXhoI obtained in Example 21 was prepared by transformation of the protoplasts of *Sac. erythraea* with the plasmid pXhoI.

The preparation of the protoplasts and the integration process were carried out as in Example 3.

The selection and the analysis of the mutants having the *ery*⁻ phenotype was carried out according to the conditions described in Example 19.

The integration into the chromosome and the presence of the expected deletion were confirmed by Southern analysis according to the general methods described in Example 19.

After transformation of the protoplasts with the plasmid pXhoI and selection of the thiostrepton resistant colonies, integration into the chromosome was confirmed by Southern analysis. Using the *Pst*I fragment of 3.3 kb of the plasmid pNCO62 as probe, the chromosomal DNA of an integrant digested beforehand with the *Pst*I and *Bgl*II restriction enzymes showed the expected 3 kb and 6 kb bands.

After repeated cultures of the integrants, the individualized colonies obtained were analyzed for sensitivity to thiostrepton and production of erythromycin, then the integration of the expected deletion (deletion of 60 bp from nucleotide 47338 to nucleotide 47397 of the sequence of Figure 3) into the chromosome of a mutant clone having the *ery*⁻ phenotype (XhoI) was confirmed by Southern analysis. The chromosomal DNA, isolated from the wild strain and the mutant XhoI respectively, was digested with the *Pst*I restriction enzyme. Hybridization with the *Pst*I-*Nco*I probe of 0.8 kb (nucleotides 46368 to 47142 of the sequence of Figure 3) produced the expected profile with a *Pst*I band of 1 kb corresponding to nucleotides 46368 to 47397 of the sequence

of Figure 3 from the wild strain and with a band of 4 kb from the mutant indicating that the *Pst*I site at position 47397 above was lost.

The loss of the *Pst*I site at position 47397 was also confirmed by PCR. The chromosomal DNA was subjected to an amplification by PCR using the primers corresponding respectively to the sequence from nucleotide 47300 to nucleotide 57320 and to the sequence from nucleotide 47661 to nucleotide 47636 of the sequence of Figure 3. An expected fragment of 306 bp was thus amplified from the wild strain generating after digestion with the *Pst*I restriction enzyme two bands of approximately 100 and 300 bp. From the mutant Xho91, a fragment of 300 bp was amplified, resulting from the deletion of 60 bp. This fragment was then isolated and was found to be resistant to digestion by the *Pst*I enzyme.

The recombinant strain thus obtained and designated Xho91, was then cultured in order to identify the metabolites produced.

EXAMPLE 23: fermentation of the Xho91 strain and identification of the secondary metabolites produced.

The culture of the Xho91strain and the analysis of the culture supernatant by mass spectrometry were carried out according to the conditions described in Example 20.

The production of erythomycin A (m/z 734 and m/z 716) was not observed but the presence of a majority quantity of erythronolide B (MK^+: m/z 441; MNa^+: m/z 425; $M-H_2O$ H^+: m/z 385) as well as the presence of desosaminyl erythronolide B (m/z 560) demonstrated characterize the Pst10 strain as an *ery*B mutant.

The mass spectrometry results were confirmed by high resolution mass spectrometry on a Brucker FT-ICR spectrometer (Brucker, FRG).

The *ery*BVI sequence has a strong homology with DnmT involved in the biosynthesis of daunorubicin in *S. peucetius* (accession number EMBL U77891) (43.9 % identity at the protein level).

These observations indicate that the *ery*BVI gene codes for the dTDP-4-keto-L-6-desoxyhexose 2,3-deshydratase

involved in the biosynthesis of dTDP-mycarose, as suggested by Scotti and Hutchinson, 1996.

EXAMPLE 24: construction of the plasmid pCIVΔ.

An integration plasmid, called pCIVΔ and carrying a deletion in the *ery*CIV gene coding for ORF17, was constructed according to the diagram in Figure 13A in the following fashion:

The plasmid pNCO62 obtained in Example 11 was digested using the *BalI* and *BclI* restriction enzymes so as to eliminate a fragment having 949 bp inside ORF17 from nucleotide 48650 to nucleotide 49598 of the sequence of Figure 3. After filling in the ends using the Klenow fragment of DNA polymerase I, the plasmid was religated and transformed in *E. coli* XL1-blue. From the plasmid pBCB17 thus generated, the fragment of 2.68 kb carrying the deletion was isolated by digestion using the *Xba*I and *Sph*I enzymes, then sub-cloned in the sites corresponding to the plasmid pUWL218. The presence of the deletion of 949 bp from nucleotide 48650 to nucleotide 49598 of the sequence of Figure 3 was confirmed by sequencing. The integration plasmid pCIVΔ thus obtained (Figure 13B) was then transferred into the *E. coli* DH5αMRC strain, then used to transform *Sac. erythraea*.

EXAMPLE 25: construction of a *Sac. erythraea* strain *ery*CIVΔ (CIV89).

A strain in which the *ery*CIV gene carries an internal deletion such as that introduced into the plasmid pCIVΔ obtained in Example 24 was prepared by transformation of the protoplasts of *Sac. erythraea* with the plasmid pCIVΔ.

The preparation of the protoplasts, the integration process and the selection of the mutants having the *ery*$^-$ phenotype were carried out as in Example 3.

In addition, the presence of the expected deletion in the chromosome (deletion of 949 bp from nucleotide 48650 to nucleotide 49598 of the sequence of Figure 3) was confirmed by genomic analysis by Southern blot as well as by PCR according to the conditions described in Example 3.

Using Southern hybridization on the genomic DNA digested

by the *Nco*I restriction enzyme, using as probe the oligonucleotide having the following sequence

C4-R AGCGGCTTGATCGTGTTGGACCAGTAC (SEQ ID No.53)

corresponding to the complementary strand of the DNA region situated from position 49996 to position 50022 of the sequence of Figure 3, a band of 6.2 kb from the wild strain and a band of 5.2 kb from the mutant CIV89 were detected. The results shown in Figure 15 indicate the presence in the mutant of a deletion of approximately 1 kb in this region of the chromosome.

The PCR amplification was carried out using the oligonucleotide having the following sequence:

C4-S GGCCTATGTGGACTACGTGTTGAACGT (SEQ ID No.54)

corresponding to the DNA region situated from position 48169 to position 48195 of the sequence of Figure 3 and the C4-R oligonucleotide having the sequence indicated above, allowing framing by PCR amplification of the region carrying the internal deletion at ORF17. Analysis by PCR amplification allowed a band of 1.8 kb in the wild strain and a band of 900 bp in the mutant CIV89 to be detected in an identical fashion to the signal obtained with the plasmid pCIVΔ. The results shown in Figure 16 confirm that the deletion of approximately 900 bp detected by Southern analysis is identical to that carried by the plasmid pCIVΔ (949 bp).

The recombinant strain thus obtained, designated CIV89, was then cultivated in order to identify the metabolites produced by the strain.

EXAMPLE 26: fermentation of the CIV89 strain and identification of the secondary metabolites produced.

The culture of the CIV89 strain and the analyses by TLC were carried out according to the conditions indicated in Example 4.

The TLC results (Figure 17) show that the CIV89 strain preferentially accumulates 3-α-mycarosyl erythronolide B as well as erythronolide B as expected of an *ery*C mutant.

Minor metabolites with low mobilities were also detected, then extracted and analyzed by RP-HPLC coupled with mass spectrometry according to the conditions used in Example

4.

A minor metabolite produces a parent peak at m/z 720 and dehydration and fragmentation products at m/z 702, m/z 576 and m/z 174. Peak 174 may correspond to 4-hydroxydesosamine and peak 576 to 4'-hydroxydesosaminyl erythronolide B.

These results suggest that the difference in m/z of 16 compared with erythromycin D (parent peak m/z 704) is carried by the aminated sugar. The proposed structure for this metabolite is 4'-hydroxy erythromycin D.

These observations indicate that the enzyme is involved in the withdrawal of the hydroxyl group in the biosynthesis route of erythromycin and that the *ery*CIV gene codes for dTDP-6-desoxyhexose 3,4-deshydratase.

The CIV89 strain was deposited at the Collection Nationale de Cultures de Microorganismes (CNCM) INSTITUT PASTEUR, 25, Rue du Docteur Roux 75724 PARIS CEDEX 15 FRANCE, on the 16 July 1997 under the number I-1905.

EXAMPLE 27: construction of the plasmid pCVΔ .

An integration plasmid, called pCVΔ and carrying a deletion in the *ery*CV gene coding for ORF18, was constructed according to the diagram in Figure 14A in the following fashion:

The *Bal*I-*Bam*HI fragment (3.48 kb), obtained from the plasmid pNCO62 prepared in Example 11 by digestion with the *Bal*I and *Bam*HI restriction enzymes, was sub-cloned in the *Sma*I-*Bam*HI sites of the vector pUC19. From the resulting plasmid pBAB18 (Figure 5B), the *Sca*I internal fragment (1kb) was then deleted by digestion with the *Sca*I enzyme in order to generate a deletion of 1044 bp from nucleotide 49998 to nucleotide 51041 of the sequence of Figure 3 in ORF18. From the plasmid pBABΔCV thus obtained, the fragment carrying the deletion was then reisolated from the polylinker of pUC19 by digestion with the *Hin*dIII and *Eco*RI restriction enzymes, then sub-cloned in the plasmid pUWL218. The integration plasmid pCVΔ thus obtained (Figure 14B) was transferred into the *E. coli* DH5αMRC strain, then used to transform *Sac. erythraea*.

EXAMPLE 28: construction of a *Sac. erythraea* strain *ery*CVΔ

(CV90).

A strain in which the *ery*CV gene carries an internal deletion such as that introduced into the plasmid pCVΔ obtained in Example 27 was prepared by transformation of the protoplasts of *Sac. erythraea* with the plasmid pCVΔ.

The preparation of the protoplasts, the integration process and the selection of the mutants having the ***ery***⁻ phenotype were carried out as in Example 3.

In addition, the presence of the expected deletion in the chromosome (deletion of 1044 bp from nucleotide 49998 to nucleotide 51041) was confirmed by genomic analysis by Southern blot as well as by PCR according to the conditions described in Example 3.

Using Southern hybridization on the genomic DNA digested by the *Nco*I restriction enzyme, using as probe the oligonucleotide having the following sequence:

C5-R AACGCCTCGTCCTGCAGCGGAGACACGAACA (SEQ ID No.55)

corresponding to the complementary strand of the DNA region DNA situated from position 51229 to position 51259 of the sequence of Figure 3, a band of 6.2 kb from the wild strain and a band of 5.1 kb from the mutant CV90 were detected. The results shown in Figure 15 indicate in the mutant the presence of a deletion of approximately 1.1 kb in this region of the chromosome.

The PCR amplification was carried out using the oligonucleotide having the following sequence:

C5-S TTCGCTCCCCGATGAACACAACTCGTA (SEQ ID No.56)

corresponding to the DNA region situated from position 49668 to position 49694 of the sequence of Figure 3 and the C5-R oligonucleotide having the sequence indicated above, allowing framing by PCR amplification of the region carrying the internal deletion at ORF18. Analysis by PCR amplification allowed a band of approximately 1.6 kb in the wild strain and a band of approximately 500 bp in the mutant CV90 to be detected in an identical fashion to the signal obtained with the plasmid pCVΔ. The results shown in Figure 16 confirm that the deletion of approximately 1.1 kb detected by the Southern analysis is identical to that carried by the plasmid

PCVΔ (1044 bp).

The recombinant strain thus obtained, designated CV90, was then cultured in order to identify the metabolites produced by the strain.

EXAMPLE 29: fermentation of the CV90 strain and identification of the secondary metabolites produced.

The culture of the strain CV90 and the analyses by TLC were carried out according to the conditions indicated in Example 4.

The TLC results (Figure 17) show that the CV90 strain preferentially accumulates 3-α-mycarosyl erythronolide B as well as erythronolide B as expected of an *ery*C mutant.

The sequence of residues 38-50 (VTGAGDGDADVQA)
Val Thr Gly Ala Gly Asp Gly Asp Ala Asp Val Gln Ala
(SEQ ID No.61)
of the protein coded by *ery*CV (sequence of SEQ ID No.11) is close to the consensus sequence binding to NAD+ described by Wierenga et al., 1985 and by Scrutton et al., 1990.

These observations indicate that the *ery*CV gene codes for a reductase which would participate as a dTDP-4,6-desoxyhexose 3,4-reductase in the biosynthesis route of d-TDP-desosamine.

EXAMPLE 30: overexpression of the product of the *ery*CIII gene in *E. coli*.

The heterologous expression of the product of the *ery*CIII gene of *Sac. erythraea* corresponding to ORF8 described in Example 1 and coding for the desosaminyltransferase activity identified in Example 7 was carried out using *E. coli* as host strain. The protein thus produced in the form of an inclusion body was then purified and its enzymatic activity determined in vitro.

1) Expression of the EryCIII protein in *E. coli*

The expression was carried out using the vector pET11a (Stratagène) for the cloning and the expression of recombinant proteins in *E. coli* under the control of the promoter of the RNA polymerase of bacteriophage T7.

Firstly, the *ery*CIII gene was amplified from the plasmid pK62 described in Example 1 in the following fashion:

The PCR amplification was carried out using Native Pfu polymerase (Stratagène) and as primers the A oligonucleotide homologous with the coding strand of the *ery*CIII gene having the sequence

A GAAGGAGATATACATATGCGCGTCGTCTTCTCCTC (SEQ ID No.57)

allowing the introduction of an *Nde*I site upstream of the initiator ATG of *ery*CIII and the B oligonucleotide homologous with the complementary strand of the *ery*CIII gene having the sequence

B CGGGATCCTCATCGTGGTTCTCTCCTTCCTGC (SEQ ID No.58)

allowing the introduction of a *Bam*HI site downstream of the stop codon of the *ery*CIII gene.

The amplified DNA was then digested by the *Nde*I and *Bam*HI restriction enzymes, then the *Nde*I-*Bam*HI fragment of 1.2 kb obtained containing the totality of the *ery*CIII gene was ligated in the expression vector pET11a (Stratagène) which contains the ampicillin-resistant β-lactamase gene, the replication origin ColE1 and the promoter of the gene of RNA polymerase T7 situated upstream of the cloning site *Nde*I, digested beforehand with the *Nde*I and *Bam*HI restriction enzymes. After ligation and transformation in *E. coli* XL1-blue, the plasmid pCEIII thus obtained was confirmed by restriction map and sequencing.

The *E. coli* BL21 strain (DE3) of the pET kit (Stratagène) which contains in its chromosomal DNA the *lac*I^q gene and the *lac*UV5 promoter upstream of the gene of RNA polymerase T7, was then transformed by the plasmid pECIII.5.

The transformed strain obtained, called BL21/pECIII, was cultured in a 50 ml erlenmeyer flask at 37°C in LB medium seeded at OD_{600}=0.1 from a preculture, then induced by isopropyl-β-D-thiogalactopyranoside (IPTG) 1mM at OD_{600}=1. After 3 hours 30 minutes of induction, 1 ml of culture was removed and centrifuged, then the bacterial pellet obtained was dissolved in 240 µl of water and 120 µl of SDS 3X sample buffer (Tris-HCl 1M pH = 6.8: 1.9 ml; glycerol 3 ml; β-mercaptoethanol 1.5 ml; SDS 20 % , 3 ml; bromophenol blue 1 % pH = 7: 0.3 ml; H_2O sqf 10 ml). From 15 µl of the solution obtained, the total proteins extracted were analyzed by SDS-

PAGE on a gel with 10 % polyacrylamide with a Comassie blue staining.

The overexpression of a protein having an apparent molecular weight of approximately 46 Kd corresponding to MW expected for the EryCIII protein was observed compared with the total proteins of a control strain transformed by the plasmid pET11a.

2) Purification of the EryCIII protein

The transformed BL21/pECIII strain above was cultured in a 6 litre fermenter in minimum medium containing glycerol as the carbon source (Korz et al., 1995) at 25°C until OD_{600}=12, then induced by IPTG for 18 hours until OD_{600}=54. From the collected broth, the bacterial pellet containing inclusion bodies was isolated by centrifugation at 5000 g for 30 mn.

Induction of the EryCIII protein was checked by SDS-PAGE (polyacrylamide gradient: 10 to 20 %) and with a Comassie blue staining after lysis on an aliquot in the 1 % SDS buffer, at 100°C for 5 min, either directly on the collected broth, or on the bacterial pellet after a first lysis by sonication in a phosphate buffer.

190 g of bacterial pellet corresponding to 1 litre of collected broth were resuspended in 2.5 volumes of KH_2PO_4/K_2HPO_4 20 mM pH 7.2 buffer containing EDTA 2.5 mM and DTT 2.5 mM. The cells were then lysed using a Rannie apparatus (Mini-Lab, type 8-30H, APV Homogenisers As, Denmark) with three passes under a pressure of 1000 bar. After centrifugation at 46.000 g for 3 hours, the pellet obtained was suspended in 2.5 volumes of 2M urea then centrifuged under the same conditions.

The pellet thus washed was then suspended in 2.5 volumes of a solution of 7M urea in a tris 50 mM pH 7.5 buffer (buffer A) so as to solubilize the EryCIII protein. After centrifugation under the same conditions, the collected supernatant obtained contains 2.1 g of total proteins determined by the Bradford method using a commercial kit (Pierce).

The extract in the 7M urea was then loaded at a rate of 0.5 metres/h and at 4-8°C on a 180 ml column (5 cm x 9 cm) of

Q sepharose (Pharmacia) equilibrated beforehand with the buffer A above and with detection at 280 nm. The EryCIII protein was then eluted with buffer A containing 0.3M NaCl. The combined fractions, containing the EryCIII protein, demonstrated by SDS-PAGE (polyacrylamide gradient: 10 to 20 %) revealed by staining with Comassie blue and 835 mg of total proteins were then loaded onto a 5.5 litre column (10 cm x 70 cm) of Superdex 200 Prep grade (Pharmacia) equilibrated beforehand with buffer A above. By elution of the column with buffer A and by detection at 280 nm, a protein peak was obtained the fractions of which containing the EryCIII protein demonstrated by SDS-PAGE and 200 mg of total proteins were combined then purified on a 180 ml column (5 cm x 9 cm) of Q Source (Pharmacia) equilibrated beforehand with buffer A. By elution with a linear gradient of NaCl varying from 0 to 0.3 M in buffer A, 30 ml of solution containing 100 mg of denatured EryCIII protein of homogeneous purity evaluated by SDS-PAGE, with revelation with silver nitrate, were obtained.

Figure 18 shows the pattern of the purity of the EryCIII protein monitored by SDS-PAGE (polyacrylamide gradient: 10 to 15 %) for a deposit of 500 ng of total proteins and revelation with silver nitrate successively after extraction with 7M urea (line 2), Q sepharose chromatography (line 3), Superdex chromatography (line 4), Q source chromatography (line 6) relative to the molecular weight markers (lines 1 and 5).

The EryCIII protein was then renatured by dilution of the homogeneous eluate with a solution of buffer A containing DTT 10 mM in order to obtain a final protein concentration of 0.1 mg/ml. The diluted solution was then dialyzed against Tris buffer 50 mM; NaCl 0.15 M; 0.3 % n-octyl-β-D-glucopyranosyl (NOG); DTT 10 mM, pH 8.3 then concentrated to 4 mg/ml by ultrafiltration on a PLGC04310 membrane (Millipore) having a cut-off threshold of 10.000.

The purified EryCIII protein was then stored in a frozen state at -20°C in aliquots of 500 µl.

3) Characterization of the EryCIII protein

The characterization of the EryCIII protein thus obtained was examined for the following properties:

a) Homogeneity.

Electrophoresis by SDS-PAGE (polyacrylamide gradient: 10 to 15 %) using the Phast System apparatus (Pharmacia) and revelation with silver nitrate shows a purity greater than 99 % for a deposit of 2000 ng.

b) Molecular weight by electrophoresis and mass spectrometry.

Using electrophoresis, an apparent MW of 46 kDa was determined in agreement with the calculated MW of 45929.

The analysis by RP-HPLC coupled with mass spectrometry (HPLC: ESI-SM) produces a mass of 45934 uma.

c) N-terminal amino acid sequence

The N-terminal sequence was determined by microsequencing on a Model A492 protein microsequencer coupled with an HPLC analyser of PTH-amino acids (Applied Biosystems).

No secondary sequence was revealed for the 10 first residues which is in agreement with the amino acid sequence described in Figure 2 (sequence of SEQ ID No.5).

d) Biological activity

The desosaminyl transferase activity of the EryCIII protein was determined in vitro by demonstration of the formation of erythromycin D from dTDP-D-desosamine, the preparation of which is described hereafter and of 3-α-mycarosyl erythronolide B (MEB) the preparation of which is described above in Materials and General Methods.

The reaction medium contains 150 nmoles of dTDP-D-desosamine, 137.4 nmoles of MEB and 1 mg of EryCIII protein using the following operating conditions:

4.78 ml of Tris buffer 50 mM pH 7.3 (buffer B); 20 µl of dTDP-D-desosamine, triethylamine salt (150 nmoles) in solution in buffer B containing EDTA 1 mM and PEFABLOC O, 4 mM (Merck); 100 µl of MEB (137.4 nmoles) in solution in buffer B and 1 mg of EryCIII protein corresponding to 250 µl of an aliquot of frozen solution obtained above are successively introduced into a glass tube provided with a screw top.

After homogenisation by Vortex, the stoppered tube is placed for 5 hours in a thermostatically-controlled bath at 30°C, then the pH is adjusted to 9-10 with 32 % NaOH then the reaction mixture is extracted 3 times with 5 ml of ethyl acetate. The extract obtained, brought to dryness under reduced pressure, then taken up in 100 µl of methylene chloride, is then analyzed by TLC under the conditions indicated in Example 4 using a methylene chloride/methanol mixture (90: 10,v/v) as eluent.

A control test (t = 0) the incubation of which is stopped immediately by the addition of NaOH, is carried out under the same conditions.

The results obtained by chemical revelation show the appearance of a less mobile product having an Rf close to the expected erythromycin D and for which a weak antibiotic activity is detected by direct autobiogram of the plates on *B. pumilus*. No biological activity is observed for the control test (Figure 19).

These results confirm that the EryCIII protein produced in *E. coli* and purified above has the expected desoaminyl transferase activity and was correctly renatured.

EXAMPLE 31: use of the sequence of the *ery*CIII gene as probe to isolate the *ole*G1 and *ole*G2 genes coding for glycosyltransferases in *S. antibioticus*.

1) cloning of the *ole*G1 and *ole*G2 genes

The sequence of the *ery*CIII gene of *Sac. erythraea* corresponding to ORF8 described in Example 1 coding for the desosaminyltransferase activity was used to prepare a hybridization probe and allowed the isolation of the homologous genes in the *S. antibioticus* ATCC 11891 strain producing oleandomycin by Southern hybridization.

The whole of the *ery*CIII gene was amplified by PCR from 6 ng of the plasmid pK62 obtained in Example 1 following the operating conditions described in Example 3 using native pfu polymerase (Stratagène) and, as primers the oligonucleotide having the following sequence:

*ery*CIII-1 CGGGTACCATGCGCGTCGTCTTCTCCTCCATG (SEQ ID No.59)

comprising a *Kpn*I restriction site in its 5' region and the

3' part of which corresponds to the complementary strand of the DNA region situated from position 10196 to position 10219 of the sequence of Figure 2 and the *ery*CIII2 oligonucleotide having the following sequence:

*ery*CIII-2 CGGGTACCTCATCGTGGTTCTCTCCTTCC (SEQ ID No.60) comprising a *Kpn*I site in its 5' region and the 3' part of which corresponds to the DNA region situated from position 8954 to position 8974 of the sequence of Figure 2.

The band of approximately 1.2 kb obtained by amplification was then digested by the *Kpn*I restriction enzyme and cloned in the plasmid pUC19 digested beforehand by the *Kpn*I enzyme. The plasmid pCIIIPCR1 thus obtained was then used to reisolate the *Kpn*I fragment of 1.2 kb corresponding to all of the *ery*CIII gene shown in Figure 2. The fragment thus isolated was then labelled with ^{32}P using the "random priming" technique described by Sambrook et al., 1989 and used as an *ery*CIII probe to analyze by Southern hybridization cosmid clones obtained from a genomic DNA library of *S. antibioticus* ATCC 11891 and prepared in the following fashion (Figure 20):

A series of six cosmids (cosAB35, cosAB76, cosAB87, cosAB67, cosAB63 and cosAB61) overlapping and covering approximately 100 kb of the region corresponding to the cluster of oleandomycin biosynthesis genes was isolated by following the method described by Swan et al., 1994 and using as probes the *Sma*I internal fragment of 2 kb at the third sub-unit of the polyketide synthase of *Sac. erythraea* in the cluster of biosynthesis genes of erythromycin (Cortes et al., 1990) followed by chromosome walking. The *str*D, *str*E and *str*M probes coding respectively for dTDP-glucose synthase, dTDP-glucose 4,6-deshydratase and dTDP-6-desoxyglucose 3,5-epimerase of *S. griseus* (Stockmann and Piepersberg, 1992) hybrize with the cosAB61 and cosAB63 cosmids (Fig. 20). In a similar fashion, by Southern hybridization with the *ery*CIII probe prepared above carried out under the standard conditions described by Hopwood et al., 1985, the cosAB35 cosmid (Swan et al., 1994) produces positive signals in two *Bam*HI restriction fragments of 3.5kb and 2.7 kb represented

in Figure 20. Subsequent sub-cloning and sequencing shows that these two fragments are separated by a *BamHI* fragment of 0.6 kb which is not detected by hybridization.

An SstI fragment of 10.8 kb of genomic DNA of *S. antibioticus* ATCC 11891 represented in Figure 21, corresponding to the right part of the cluster of oleandomycin biosynthesis genes comprised between the SstI site in position 11081 of the OLE-ORF3 gene of the PKSs of the EMBL No.L09654 sequence) and the SstI site in position 5 of the EMBL No.L36601 sequence situated 1.4 kb upstream of the oleB gene and hybridizing with the *ery*CIII probe prepared above, was starting isolated from the cosAB35 cosmid and sub-cloned in the plasmid vector pSL1180 (Pharmacia Biotech). The pCO35-S clone thus obtained was used to generate single-stranded templates by sub-cloning different DNA fragments in the M13mp18 and MP13mp19 bacteriophages (New England Biolabs), then the nucleotide sequence of these fragments was determined according to the method of Sanger et al. (1977) using a modified T7 polymerase (Sequenase version 2.0; U.S. Biochemicals) in the presence of $\alpha[^{35}S]$dCTP (Amersham) and 7-deaza-dGTP, according to the supplier's recommendations in order to limit band compression problems. The standard primers supplied with the Sequenase kit as well as the internal synthetic primers(17mer) were used.

Assembly of the sequence data was carried out using the Fragment Assembly program (Genetic Computer Group, University of Wisconsin) and identification of the open reading frames using the CODONPREFERENCE program (Devereux et al., 1984).

The nucleotide sequences obtained allowed the nucleotide sequence of 6093 bp represented in Figure 22(sequence of SEQ ID No.15) to be established, comprised between the *Sph*I* and *Kpn*I sites shown in Figure 21, in which five ORFs were identified from nucleotide 184 to nucleotide 1386 (ORF called *ole*P1), from nucleotide 1437 to nucleotide 2714 (ORF called *ole*G1), from nucleotide 2722 to nucleotide 3999 (ORF called *ole*G2), from nucleotide 3992 to nucleotide 4729 (ORF called *ole*M) and from nucleotide 4810 to nucleotide 5967 (ORF called *ole*Y) respectively). The five ORFs are transcribed in the

same direction.

Specimens of *E. coli* containing the plasmid pCO35-S comprising the coding region of ORFs *ole*P1, *ole*G1, *ole*G2, *ole*M and *ole*Y were deposited at the Collection Nationale de Cultures de Microorganismes (CNCM) INSTITUT PASTEUR, 25, Rue du Docteur Roux 75724 PARIS CEDEX 15 FRANCE, on the 8 April 1998 under the number I-2003.

The *ole*G1 gene codes for a polypeptide having 426 amino acids (sequence of SEQ ID No.17). However, the presence of a CGC codon coding for a highly preserved arginine in this class of glycosyltransferase in the Streptomyces situated immediately upstream of the GTG codon, would indicate that the initiation codon could be the CTG codon in position 1431 of the sequence SEQ ID No.17.

The *ole*G2 gene codes for a polypeptide having 426 amino acids (sequence of SEQ ID No.18).

Comparison of the deduced amino acid sequences of the *ole*G1 and *ole*G2 ORFs above with the database proteins using the Blast program (Altschul et al., 1990) showed similarities with glycosyl transferases from different sources, in particular approximately 72 % similarity and 53 % identity with the EryCIII deosaminyltransferase described in Example 30.

Identification of the function of the *ole*G1 gene or the *ole*G2 gene was then carried out by interruption of the target gene in the *S. antibioticus* strain ATCC 11891 and by identification of a non-glycosylated precursor of the oleandomycin produced by the mutant strain using the methods described in Examples 3 to 4.

2) generation of an *S. antibioticus oleG1Δ* strain (A35G1).

A strain in which the *ole*G1 gene is interrupted was prepared by integration of a plasmid pCO3 into the homologous regions of the chromosomal DNA of the of *S. antibioticus* strain ATCC 11891 producting oleandomycin.

Firstly, the *Bam*HI fragment of 0.6 kb internal to the *ole*G1 gene, obtained by digestion of the plasmid pCO35-S prepared above, with the *Bam*HI restriction enzyme (Figure 21), was sub-cloned in the *Bam*HI site of the plasmid pOJ260

NRRL B-14785.

The plasmid pCO3 thus generated was then transferred into the *E. coli* strain TG1 *rec*O1504::Tn5 (Kolodner et al., 1985), then used to transform the protoplasts of *S. antibioticus*. Selection of the transformants was carried out by resistance to apramycin (injection-grade Apramycin, Rhône Mérieux).

Preparation of the protoplasts was carried out from the *S. antibioticus* strain ATCC 11891 following the conditions described by Hopwood et al., 1985.

Transformation was carried out using 50 µl of an aliquot of protoplasts, 5 µg of plasmidic DNA pCO3 and replacing the thiostrepton with apramycin at the final concentration of 25 µg/ml.

Selection of the integrants carried out by resistance to apramycin allowed a clone called A35G1 to be isolated.

The expected alteration in the corresponding region of the chromosome of *S. antibioticus* was confirmed by genomic analysis by Southern blot. The chromosomal DNA isolated, then digested by the *Pst*I restriction enzyme from the wild *S. antibioticus* strain or the A35G1 mutant was analyzed by Southern using the *Bam*HI fragment of 0.6 kb indicated above as hybridization probe. Replacement of the *Pst*I fragment of 4.7 kb thus detected in the wild strain by two *Pst*I fragments of 2.4 and 6.5 kb in the A35G1 mutant confirms the integration of the plasmid pCO3 into the chromosome of the A35G1 strain at the level of *ole*G1 ORF.

The A35G1 recombinant strain thus obtained was then cultured to identify the precursors produced by the strain.

3) fermentation of the A35G1 strain and identification of the oleandomycin precursors produced.

The A35G1 strain was cultured for 72 hours in a 50 ml erlenmeyer in EP2 medium from a 48 hour preculture in an EP1 medium under the conditions described in Example 4.

The broth extracts with ethyl acetate were then analyzed according to the methods used in Examples 3 and 4.

a) The biological test by record of bacterial sensitivity to antibiotics was carried out in the following

fashion:

After growth of the *B. pumilus* strain on TSB medium overnight at 37°C, the culture was diluted to 1/100 in medium containing 50 % (w/v) of glycerol, then the cell suspension obtained was preserved at -20°C before use.

The biological test was then carried out by introducing 150 µl of the thawed cell suspension into 100 ml of TSB medium containing 1 % agar and kept at 55°C. The mixture was then poured into petri dishes. After cooling down, Oxford cylinders containing 50 to 200 µl of ethyl acetate extracts were placed on agar dishes, kept for 2 hours at 4°C, then incubated overnight at 37°C.

The extracts show no inhibitory effect on the growth of *B. pumilus* ATCC 14884.

b) The analysis by TLC by chemical revelation was carried out according to the conditions described in Example 4 using as standards erythromycin A, erythronolide B as well as 6-desoxyerythronolide B.

The analysis by TLC shows that the A35G1 strain does not produce oleandomycin but preferentially accumulates a crimson product having a mobility which is greater than erythronolide B and close to 6-desoxyerythronolide B and that the aglycone 8,8a-desoxyoleandolide part can be expected.

c) The analysis by RP-HPLC chromatography coupled with mass spectrometry was carried out according to the conditions described in Example 4. Two major metabolites, called M9 and M10, were detected (elution at 6.12 mn and 17.23 mn respectively). The two products produce a parent peak at m/z 373 and similar fragmentation profiles which may be in accordance with the structure [8,8a-desoxyoleandolide]H^+. However, only the retention time of the M10 metabolite is in accordance with the proposed structure whilst the M9 metabolite could correspond to an isomeric structure or to the open lactone nucleus.

Experiments complementing the mutant strains of *Sac. erythraea* CIII68 described in Example 6 and BV88 described in Example 16 were also carried out using plasmidic constructions allowing respectively each of the *ole*G1 and

*ole*G2 genes to be expressed.

These observations and the absence of detection of oleandrosyl 8,8a-desoxyoleandolide or desosaminyl 8,8a-desoxyoleandolide indicate that the *ole*G1 gene codes for the desoaminosyltransferase and the *ole*G2 gene codes for the oleandrosyltransferase respectively involved in the biosynthesis of oleandomycin.

PREPARATION OF EXAMPLE 30: Thymidine 5'-(trihydrogen-diphosphate),P'-[3.4,6-trideoxy-3-(dimethylamino)-D-xylo-hexopyranosyl]ester,N,N-diethylethanamine

STAGE A: 3,4,6-trideoxy-3-(dimethylamino)-D-xylohexopyranose hydrochloride

146.6 g of erythromycin A is added under agitation and at ambient temperature to 1.5 litres of 6N hydrochloric acid The solution obtained is taken to reflux for 2 hours. The reaction medium is cooled down to ambient temperature, filtered and the residue obtained is washed with water. The aqueous phase is extracted with methylene chloride, then with sulphuric ether. 10 g of black L_2S is added to the aqueous phase and the traces of ether are driven off under reduced pressure. After filtering and concentrating, a second run is carried out under the same conditions. The two tests are combined, dissolved in ethanol (150 cm^3), 150 cm^3 of ethyl ether is added. The crystals obtained are separated, washed and dried. 42 g of sought product is obtained. M.p. = 158~160°C.

STAGE B: 3,4,6-trideoxy-3-(dimethylamino)-D-xylo-hexopyranose,1,2-diacetate

60 cm^3 of triethylamine is added under agitation at 20°C, to a mixture containing 15.27 g of the product of Stage A and 150 cm^3 of methylene chloride. A solution containing 20 cm^3 of acetic anhydride and 80 cm^3 of methylene chloride is added at 20°C. Agitation is carried out at ambient temperature for 20 hours, followed by filtering, washing, concentrating and impasting in sulphuric ether. The filtrate is concentrated under reduced pressure. The residue obtained is chromatographed eluting with an ethyl acetate-triethylamine mixture (95-5). 18.6 g of sought product is obtained which

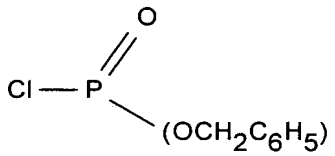
is used as such in the following stage.

STAGE C: 3,4,6-trideoxy-3-(dimethylamino)-D-xylo-hexopyranose,2-acetate

A mixture containing 18.6 g of the product of Stage B and 50 cm^3 of DMF is taken to 50°C and 6.62 g of hydrazine acetate NH_2NH_2, ACOH is added. The reaction mixture is agitated and poured into a saturated solution of sodium acid carbonate. The aqueous phase is extracted with ethyl acetate. The organic phases are combined, dried, filtered and concentrated, followed by distillation under reduced pressure in order to eliminate the DMF by azeotropic entrainment with toluene. 11.28 g of product is obtained which is chromatographed on silica eluting with a ethyl acetate-triethylamine mixture (90-10). 6.5 g of sought product is obtained which is used as such in the following stage.

STAGE D: 3,4,6-trideoxy-3-(dimethylamino)-D-xylo-hexopyranose,2-acetate bis(phenylmethyl)phosphate

5.7 cm^3 of a solution of n-butyllithium in hexane is added at -70°C to a solution containing 1.738 g of the product of the preceding stage and 40 cm^3 of THF. 10 g of dibenzylphosphochloride prepared extemporaneously (J. Chem. Soc. 1958, p. 1957),



in solution in 20 cm^3 of THF is added at -70~-75°C. Agitation is maintained for 1 hour 30 minutes between -70 and -74°C. The reaction medium is poured into a saturated solution of sodium acid carbonate and extracted with ethyl acetate. The organic phase is dried over sodium sulphate, filtered and concentrated. The product obtained is chromatographed on silica eluting with an acetone-methylene chloride mixture (5-5). 1.070 g of sought product is obtained.

STAGE E: 3,4,6-trideoxy-3-(dimethylamino)-D-xylo-hexopyranose,1-(dihydrogen phosphate),N,N-diethylethanamine

A mixture containing 1.070 g of the product of the preceding stage, 20 cm^3 of ethyl acetate, 10 cm^3 of methanol, 0.622 cm^3 of triethylamine and 200 mg of palladium on carbon is placed under agitation and under a hydrogen flow for 30 minutes at ambient temperature. The reaction medium is filtered, washed with methanol and with ethyl acetate and the filtrate is concentrated. 1 g of an oil is obtained to which 10 cm^3 of methanol is added. The solution obtained is agitated for 20 hours. The methanol is driven off under reduced pressure at 30°C. 680 g of sought product is obtained.

STAGE F: 3,4,6-trideoxy-3-(dimethylamino)-D-xylo-hexopyranose,1-(dihydrogen phosphate)

420 mg of the product isolated in the form of triethylamine salts, obtained in the preceding stage is dissolved in 1.6 cm^3 of methanol. 3.2 cm^3 of sulphuric ether is added, then 6.4 cm^3 of sulphuric ether is added. 250 mg of sought product is obtained melting at 242~244°C.

STAGE G: Thymidine 5'-(trihydrogen diphosphate),P'-[3.4,6-trideoxy-3-(dimethylamino)-D-xylo-hexopyranosyl]ester,N,N-diethylethanamine

228 mg of the product of Preparation 1, 6 cm^3 of pyridine and 544 mg of thymidine 5'-monophosphate morpholidate-4-morpholine-NN'-dicyclohexylcarboxamidine are mixed together. The pyridine is driven off under reduced pressure using a rotary evaporator while maintaining the temperature at 30°C or below. 6 cm^3 of pyridine is added which is driven off under reduced pressure. The operation is repeated twice. 6 cm^3 of pyridine and 105 mg of 1H-tetrazole are added. Agitation is carried out for 3 days at ambient temperature. The pyridine is driven off under reduced pressure, followed by taking up in water, filtering, concentrating and a product is obtained which is purified by chromatography. In this way the sought product is obtained.

rf = 0.12 eluent CH_2Cl_2, MeOH, H_2O (60-35-6).

Bibliographical references:

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) J Mol Biol 215: 403-410.

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA Struhl K (1995) Current protocols in molecular biology. Massachusetts General Hospital and Harvard medical School. John Wiley and Sons, Inc.

Bankier AT, Weston KM, Barrell BG (1987) Methods in Enzymology 155: 51-93.

Bevitt DJ, Cortes J, Haydock SF and Leadlay PF, (1992), Eur. J. Biochem 204: 39-49.

Caffrey P, Bevitt DJ, Staunton J, Leadlay PF (1992) FEBS 304: 225-228.

Cortés J, Haydock SF, Roberts GA, Bevitt DJ, Leadlay PF (1990) Nature 348: 176-178.

Devereux J, Haeberli P, Smithies O (1984) Nucl Acids Res 12: 387-395.

Dhillon N, Hale RS, Cortes J, Leadlay PF (1989) Mol Microbiol 3: 1405-1414.

Dickens ML, Ye J, Strohl WR (1996) J Bacteriol 178: 3384-3388.

Donadio S, Stassi J, McAlpine JB, Staver MJ, Sheldon PJ, Jackson M, Swanson SJ, Wendt-Pienkowski E, Wang YG, Jarvis B, Hutchinson CR, Katz L (1993) In: Baltz RH, Hegeman GD, Skatrud PL (eds) Industrial Microorganisms: Basic and Applied Molecular Genetics: American Society for Microbiology, Washington, DC, 257-265.

Gandecha AR, Large SL, Cundliffe E (1997) Gene 184: 197-203.

Haydock SF, Dowson JA, Dhillon N, Roberts GA, Cortes J, Leadlay PF (1991) Mol Gen Genet 230: 120-128.

Hessler PE, Larsen PE, Constantinou AI, Schram KH, Weber JM, Appl Microbiol. Biotechnol (1997), 47: 398-404.

Hopwood DA, Kieser T, Wright HM an Bibb MJ, Journal of General Microbiology (1983), 129, 2257-2269.

Hopwood DA, Bibb MJ, Chater KF, Kieser T, Bruton C, Kieser HM, Lydiate DJ, Smith CP, Ward JM, Schremp H (1985) Genetic manipulation of Streptomyces. A laboratory Manual. John Innes Foundation, Norwich.

Kaneda T, Butte JC, Taubman B, Corcoran JW (1962) J Biol Chem 237: 322-327.

Katz E, Thompson CJ, Hopwood DA (1983) J Gen Microbiol 129: 2703-2714.

Korz DJ, Rinas U, Hellmuth K, Sanders EA, Deckwer W-D (1995) Journal of Biotechnology 39: 59-65.

Katz L, Donadio S (1995) Macrolides. In: Vining LC, Stuttard C (eds). Genetics and Biochemistry of Antibiotic Production. Butterworth-Heinemann, Newton, MA, 385-420.

Liu H-W, Thorson JS (1994) Annu Rev Microbiol 48: 223-56.

Otten SL, Liu X, Ferguson J, Hutchinson CR (1995) J Bacteriol 177: 6688-6692.

Sakakibara H and Omura S (1984). In: Omura S. (ed) Macrolide Antibiotics: Chemistry, Biology, and Practice. Academic Press, Inc. London, 85-125.

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, N.Y.

Sanger F, Nicklen S, Coulson AR (1977) Proc Natl Acad Sci USA 74: 5463-5467.

Scotti C, Hutchinson CR (1996) J Bacteriol 178: 7316-7321.

Scrutton NS, Berry A, Perham RN (1990) Nature 343: 38-43.

Staden R (1984), Nucleic Acids Res 12: 521-528.

Stassi D, Donadio S, Staver MJ, Katz L (1993) J Bacteriol 175: 182-189.

Stemmer WPC (1994) Proc. Natl. Acad. Sci., Vol. 91, pp. 10747-10751.

Stockmann M, and Piepersberg W (1992) FEMS Microbiol Lett 90: 185-190.

Swan D.G., Rodriguez A.M., Vilches C., Méndez C., Salas J.A., (1994) Mol Gen Genet 242: 358-362.

Ward JM, Janssen GR, Kieser T, Bibb MJ, Buttner MJ, Bibb MJ (1986) Mol. Gen. Genet. 203: 468-475

Weber JM, Wierman CK, Hutchinson CR (1985) J Bacteriol 164: 425-433.

Weber JM, Losick R (1988) Gene 68: 173-180.

Weber JM, Schoner B, Losick R (1989) Gene 75: 235-241.

Weber JM, Leung JO, Maine GT, Potenz RHB, Paulus TJ, DeWitt JP (1990) J Bacteriol 172: 2372-2383.

Weber JM, Leung JO, Swanson SJ, Idler KB, McAlpine JB (1991) Science 252: 114-117.

Wehmeier UF (1995) Gene 165: 149-150.

Wierenga RK, Terpstra P, Hol WGJ (1986) J. Mol. Biol. 187: 101-107.

Yamamoto H, Maurer KH, Hutchinson CR (1986) J Antibiot 34: 1304-1313.

Free text in the sequence listings

SEQ ID No.: 1:
/function= "involved in the biosynthesis of mycarose"
/gene= "eryBII"
/function= "involved in the biosynthesis of desosamine"
/gene= "eryCII"

SEQ ID No.: 4:
/function= "involved in the biosynthesis of desosamine"
/gene= "eryCIII"
/note= "SEQ ID No 1 FROM 1046 TO 2308"

SEQ ID No.: 6:
/function= "involved in the biosynthesis of mycarose"
/gene= "eryBIV"
/transl_except= (pos: 242 .. 244, aa: Met)
/function= "involved in the biosynthesis of mycarose"
/gene= "eryBV"
/transl_except= (pos: 1210 .. 1212, aa: Met)
/function= "involved in the biosynthesis of desosamine"
/gene= "eryCVI"
/function= "involved in the biosynthesis of mycarose"
/gene= "eryBVI"
/transl_except= (pos: 3308 .. 3310, aa: Met)
/function= "involved in the biosynthesis of desosamine"
/gene= "eryCV"
/function= "involved in the biosynthesis of mycarose"
/gene= "eryBVII"
/transl_except= (pos: 7578 .. 7580, aa: Met)

SEQ ID No.: 13:
/function= "involved in the biosynthesis of desosamine"
/gene= "eryCIV"
/note= "SEQ ID No 6 FROM 4837 TO 6039"

SEQ ID No.: 15:
/gene= "oleP1"

/function= "glycosylation of 8,8a-desoxyoleandolide"
/gene= "oleG1"
/transl_except= (pos: 1437 .. 1439, aa: Met)
/function= "glycosylation of 8,8a-desoxyoleandolide"
/gene= "oleG2"
/gene= "oleY"

SEQ ID No.: 20:
/gene= "oleM"
/note= "SEQ ID No 15 FROM 3992 TO 4729"

SEQ ID No.: 22 to SEQ ID No.: 60
/desc = "OLIGONUCLEOTIDE"

SEQ ID No.: 61:
/note= "SEQ ID No 11 FROM 38 TO 50"